TRACKING BRAIN DYNAMICS ACROSS TRANSITIONS OF CONSCIOUSNESS

IULIA-MARIA COMȘA

LUCY CAVENDISH COLLEGE

DEPARTMENT OF CLINICAL NEUROSCIENCES

UNIVERSITY OF CAMBRIDGE

This dissertation is submitted for the degree of Doctor of Philosophy.

APRIL 2018

ABSTRACT

How do we lose and regain consciousness? The space between healthy wakefulness and unconsciousness encompasses a series of gradual and rapid changes in brain activity. In this thesis, I investigate computational measures applicable to the electroencephalogram to quantify the loss and recovery of consciousness from the perspective of modern theoretical frameworks. I examine three different transitions of consciousness caused by natural, pharmacological and pathological factors: sleep, sedation and coma.

First, I investigate the neural dynamics of falling asleep. By combining the established methods of phase-lag brain connectivity and EEG microstates in a group of healthy subjects, a unique microstate is identified, whose increased duration predicts behavioural unresponsiveness to auditory stimuli during drowsiness. This microstate also uniquely captures an increase in frontoparietal theta connectivity, a putative marker of the loss of consciousness prior to sleep onset.

I next examine the loss of behavioural responsiveness in healthy subjects undergoing mild and moderate sedation. The Lempel-Ziv compression algorithm is employed to compute signal complexity and symbolic mutual information to assess information integration. An intriguing dissociation between responsiveness and drug level in blood during sedation is revealed: responsiveness is best predicted by the temporal complexity of the signal at single-channel and low-frequency integration, whereas drug level is best predicted by the complexity of spatial patterns and high-frequency integration.

Finally, I investigate brain connectivity in the overnight EEG recordings of a group of patients in acute coma. Graph theory is applied on alpha, theta and delta networks to find that increased variability in delta network integration early after injury predicts the eventual coma recovery score. A case study is also described where the re-emergence of frontoparietal connectivity predicted a full recovery long before behavioural improvement.

The findings of this thesis inform prospective clinical applications for tracking states of consciousness and advance our understanding of the slow and fast brain dynamics underlying its transitions. Collating these findings under a common theoretical framework, I argue that the diversity of dynamical states, in particular in temporal domain, and information integration across brain networks are fundamental in sustaining consciousness.

ACKNOWLEDGEMENTS

I owe my gratitude to those who kindly offered me guidance and inspiration in my daily life as a doctoral student at Cambridge.

I am grateful to Dr. Srivas Chennu for supervising me as closely and meticulously as I could have ever hoped for, while giving me the freedom to decide on my own research explorations; for always being there to help with everything I ever needed to make progress, while gently but steadily pushing me forward. I am also grateful to Dr. Tristan Bekinschtein for co-supervising me and for welcoming me into his lab, a charming world of wonders that reflects his ingenious mind and that fostered both my intellect and my soul. I thank Prof. Anil Seth and Dr. Emmanuel Stamatakis for kindly agreeing to examine my thesis and for a thorough and stimulating discussion about consciousness and its transitions. Finally, I hold particular gratitude to Prof. Joe Herbert for placing his trust in me when I applied to study neuroscience at Cambridge; his support has been, and still is, inspiring me to do my best.

I cheerfully thank my lab companions: Andrés Canales Johnson, for discussions full of insight, for brilliantly sarcastic humour and for his knowledgeable and outspoken opinions on all aspects of consciousness; Anat Arzi, for her supportive advice about being a research student, for her always thoughtful questions and for brightening the lab with her lovely spirit; Sri Jagannathan, for discussions on both advanced methodologies and on the culture and mysteries of the human race; Maria Niedernhuber, Alejandro Ezquerro Nassar, Barbara Jachs, and everyone else in the lab for creating a most pleasant research environment; and Valdas Noreika, for allowing me to inherit the workspace of a researcher with a warm and gentle temperament, whose work I greatly admire.

I am grateful to the Cambridge University Karate Club for strengthening my spirit and lighting up the dark days when my research work was difficult. Progressing from clumsily entering its dojo to competing for Cambridge and earning trophies, I unexpectedly found a new passion that taught me I must care for my body as much as for my mind.

I kindly thank everyone who graced me with friendship and love during my years at Cambridge; old and new, family and friends, from far away or from close by; offering words of support, poems and songs to dream by, or intriguing late-night conversations about the nature of reality; I am most lucky for these moments that beautifully shaped my mindscape.

DEDICATION

First and foremost, to my parents: my kind, dedicated, compassionate mother, and my determined, diligent, playfully creative father. I love you more than anything in the world. Thank you for giving me life, and for your unconditional love. My deepest wish is for you both to live forever.

Secondly, to my teachers from secondary school, who nourished my curiosity, my intellect, my creativity, and my evolution as a human being. To Dorina Zaharia, my principal teacher, who taught mathematics with dedication, intellectual elegance and appropriate severity, and guided us with wisdom during the delicate age between childhood and adolescence. To Sybille Balint, my teacher of Romanian literature and grammar, who encouraged us to follow our dreams, incited us to find our true passion and empowered us to express ourselves freely in a world of constraints. To all of my other teachers from General School No. 8 of Brașov, who nurtured my growth, and to my classmates, beautiful souls, as rebellious as they were intelligent; I would not be myself today had it not been for those four years of my life in this school.

Finally, to dreams, and to the endless pursuit of understanding. For me, it began when I was 14, one early morning, as I was still half-asleep sitting on my chair in the kitchen. I closed my eyes and tried to observe the shimmering light of awareness that arose in my mind as I was waking up, before any other thoughts appeared. What I discovered was a wonderous new inner state of pure consciousness, which I felt must hold the essence of one's existence. I think this was the moment when I decided that, to me, consciousness is the most worthwhile thing that I can ever hope to understand. For the next fifteen years, I have been doing my best to safeguard my inner child's sense of wonder and curiosity, and the piercing desire to understand what lies at the boundary of wakefulness. I have been exploring fascinating inner worlds of hypnagogia, lucid dreams and synesthesia; and I have trained my mind, to the best of my ability, to keep all of the awe under the scrutiny of critical thinking and to explore it in computational ways. Neuroscience is one avenue I also chose to follow, and so this thesis humbly holds one attempt, employing one of the many sets of tools available to the mind, to progress in the quest for understanding consciousness. I hope that many more will follow and that the pursuit of knowledge will always be a guiding light for humankind.

"This message is for whatever entities will come to be after this world is gone. It might not be necessary. You might know everything that exists in my mind. I know that to you, with access to information I can only dream of, my theories must appear ridiculous. I barely understand anything of the world I inhabit. I want you to know that I understood that, but tried anyway. It matters that an objective truth exists and that we struggle to understand it."

- The Talos Principle

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the Clinical Medicine and Clinical Veterinary Medicine Degree Committee.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
AUC	Area Under the Curve
BOLD	Blood-Oxygen-Level Dependent
CRS-R	Coma Recovery Scale-Revised
DoC	Disorder of Consciousness
EEG	Electroencephalogram
fMRI	functional Magnetic Resonance Imaging
GEV	Global Explained Variance
GFP	Global Field Power
GCS	Glasgow Coma Score
LZ	Lempel-Ziv (also see Table 4.1)
LZW	Lempel-Ziv-Welch
ICA	Independent Component Analysis
MEG	Magnetoencephalography
PET	Positron Emission Tomography
PLE	Phase Lag Entropy
PLI	Phase Lag Index
PLV	Phase Locking Value
ROC	Receiver Operating Characteristic
RT	Reaction Time
SVM	Support Vector Machine
SD	Standard Deviation
TMS	Transcranial Magnetic Stimulation
WPLI	Weighted Phase Lag Index
WSMI	Weighted Symbolic Mutual Information

CHAPTER 1

CONSCIOUSNESS

This chapter introduces the main theme of this thesis: consciousness. For millennia, consciousness has belonged to the domain of philosophy. Over the last century, inheriting ideas and question from great thinkers of the past, scientists have also begun to study this topic. This chapter starts by telling the story of why consciousness is one of the most fascinating topics among all of those ever known to humankind. Nonetheless, defining the concept of consciousness is a great challenge; the chapter continues by presenting several ideas and thought experiments that have been historically proposed in relation to consciousness. Then, the chapter focuses on the experimental topic of this thesis and describes why sleep, anaesthesia, sedation, and disorders of consciousness are of uttermost interest within this field, while also clarifying the theoretical limits of what current neuroscientific approaches can achieve. Finally, several theories of consciousness that have recently emerged as valuable, both conceptually and due to their practical applications, are presented with their advantages and disadvantages, concluding with a clear theoretical framework for the neurocomputational investigations in the rest of this thesis.

1.1. THE RIDDLE

Somewhere in the unimaginable vastness of the universe (Sokolov and Shvartsman, 1975), on a planet orbiting a star located in a minor spiral arm of the Milky Way galaxy (Hawking, 2001), life emerged and evolved (Darwin, 1859; Oro et al., 1990) into elegant mesoscale beings that could perceive the world through multiple senses, adapt to and change their own environment, feel emotions, communicate, and even entertain complicated internal monologues about their own existence (Descartes, 1641). Once these creatures evolved enough that fighting for survival was no longer a daily concern, they pursued more abstract ideas and goals, creating arts, religions, literature, philosophy and sciences, and continuously pushing the boundaries of their knowledge (Bryson, 2003). They even created artificial computational entities, which one day, endowed with finely-engineered intelligence (Bachman et al., 2016), might be able to read and summarise this very chapter.

If such an artificially intelligent reader were to read this account, it might be puzzled to find out that many humans came to the conclusion that there was something seemingly ineffable about their own

existence: they had a first-person, subjective perspective of the world (Nagel, 1974). They tried, for example, to explain everything about the colour they called 'red', and although they knew precisely how to describe in nanometres the exact wavelength of light that caused this percept and the exact pathway in their brains that eventually led to the uttering of the word 'red', many felt like something was missing from this comprehensive, functionally perfect description (Chalmers, 1995). These people claimed that the *experience* of seeing this colour, despite being at the heart of this question, was simply left aside from its otherwise wholesome answer, and wondered why.

Over the years, this abstract problem took many forms. For a long time, many humans thought that perceptual, emotional, or cognitive capacities resided in immortal souls, while their bodies were only perishable vessels; most cultures around the world held a variation of such beliefs (Eliade, 1985, 1982, 1981). Throughout history, some philosophers identified the soul with the mind (Descartes, 1641), some considered the soul and the spirit as separate entities (Aristotle, 350BC), and some simply discarded any separation between the two and the body (Dennett, 1993). Dualism has been the most withstanding position in philosophy of mind, and has continued to persist among the general and educated public (Demertzi et al., 2009), as well as scientists (Larson and Witham, 1997).

Consciousness seemed to be intimately linked with other cognitive functions, such as attention (Koch and Tsuchiya, 2007), memory (Clark and Squire, 1998), emotion (Damasio, 1999) and intelligence (Århem and Liljenström, 1997). These cognitive functions were more straightforward to study with typical approaches in psychology and neuroscience. It took a while until consciousness became accepted as a valid topic of investigation by science (Crick and Koch, 1990). When this happened, the focus of the quest became the reconciliation of the phenomenological first-person view, most commonly judged through behavioural events or verbal reports, with markers of brain activity at microscopic and macroscopic levels (Crick and Koch, 2003). Many dimensions arose in the landscape of questions about consciousness: sleep versus wakefulness, wakefulness versus awareness, reportability versus phenomenology. A rich body of evidence accumulated to support certain hypotheses, as will be detailed in the next sections, and lively academic debate ensued.

Whether this is only the beginning of the story or an outcome is imminent, research on the topic of consciousness is currently more dynamic and fast-paced than it has ever been, largely because of recent developments in brain imaging techniques, but also due to conceptual advances. In this context, this thesis aims to contribute to the story of consciousness by examining the behavioural loss of consciousness during the transitions to sleep and sedation, and the restoration of consciousness after coma.

Despite a number of radically different views throughout history, as well as recent debate sparkled by the development of artificial intelligence, a human brain currently seems the only entity in the universe that we can be reasonably certain sustains consciousness. Therefore, in the next section, for the benefit of those interested in consciousness who might not possess a background in biology, a summary of what we know about the brain will be given, emphasising the relationship between its anatomy and its functions, as well as the challenge that this relationship poses to the quest of understanding consciousness.

1.2. THE HUMAN BRAIN

It has been known since ancient times that the brain is where all emotion, cognition, movement and mental wellbeing or disease originate (Hippocrates, 400BC). The human brain is the centre of the nervous system and consists of neurons and glial cells. Neurons are believed to be the main units in the nervous system, which respond to sensory input, command motor responses, and carry out various information by processing tasks at lower and higher levels. Glia is thought to be mainly responsible for providing a nourishing environment for neurons to carry their tasks, although evidence has been building up that they also participate in neurotransmission (Allen and Barres, 2009). Human brains have around 10^{11} neurons and around 10 times more glial cells (Bear et al., 1996).

Neurons are cells with a diameter of 0.01 to 0.05 mm typically consisting of a main body called soma, an axon which can extend for as long as a meter which transmits information to other potentially long-distance neurons, and several dendrites which form a tree-like structure with branches gathering information transmitted by axons belonging to other neurons. Communication between neurons is achieved through action potentials, impulses transmitted as electrical charges along the membrane of the axon. The point of information exchange between axons and dendrites, called a synapse, involves the release of neurotransmitters with either excitatory or inhibitory effect. A typical neuron has on order of 10^3 synapses, although some types of neurons may have around 10^5 synapses (Gazzaniga et al., 2014). Computationally, a neuron can hence be described as a nonlinear (Larkum and Nevian, 2008) processing unit with inputs and outputs. The frequency of action potentials of individual neurons, as well as the distribution of action potentials within a population of neurons, generates information (Bear et al., 1996).

While understanding the cellular fundamentals of how the brain works is important, higher brain functions like cognition need to be ultimately explored by looking at the higher level of neural organisation, circuits, systems and brain areas, which work together anatomically and functionally. The functions performed by the human brain are to some extent reflected in its anatomical

organisation (Gazzaniga et al., 2014). Many life-supporting functions that do not require consciousness are performed by the structures of the brainstem: the medulla, which supports respiration, heart rate and arousal; the pons, which modulates both arousal and rapid eye movement (REM) sleep; the cerebellum, which, astonishingly, houses the majority of neurons in the central nervous system, and which integrates information about the body and motor commands, in order to modulate fine motor coordination and other higher cognitive functions; and the midbrain, which plays a role in gaze and motor coordination. Above the brainstem sits the thalamus, which acts as a gateway to the cortex, as it receives and passes on information received from all sensory modalities, except for olfaction, through specialised nuclei. The hypothalamus is responsible for homeostasis; it initiates feelings such as hunger or thirst through the endocrine system, and controls circadian cycles. The basal ganglia consist of several structures that are thought to be essential in goal-oriented and reward-based behaviour and learning, with dopamine possibly signalling the prediction error of the reward.

As we begin to investigate higher cognitive functions, it becomes more challenging to unify neural anatomy and function. The evolutionarily newest part of the central nervous system is the cerebral cortex, which consists of layers of cells densely packed and folded within the space constraints of the skull. Although it has been tentatively partitioned into several anatomical configurations using histological analyses to detect similar and dissimilar areas, the specific functions of each particular area can be difficult to pin down. Visual information is processed, in a hierarchical manner, in the visual cortex. Auditory information is processed in the temporal lobe, although this lobe mediates other functions as well, such as memory, emotion and language comprehension. The parietal lobe is responsible, among other things, for attention and spatial reasoning. The somatosensory and motor cortices are organised into a topographical map of the body, that receive inputs and send commands respectively from and to the body. Finally, the prefrontal cortex is involved in higher cognitive functions like planning, organising and executing actions.

To summarise, our conscious experience seems to emerge from an intricate interplay of cognitive functions associated with multiple brain areas, as opposed to a single location in the brain. To understand consciousness, we need to look at emergent characteristics of brain function. Moreover, we should consider that information processing in the brain might be performed beyond algorithmically operating with local representations of environmental inputs (Williams, 2018). One alternative account is that of the brain as a prediction machine that continuously makes guesses and updates its model of the world, which might offer a better conceptual framework for understanding consciousness (Seth, 2016).

1.3. DEFINING CONSCIOUSNESS

Although questions about consciousness have evolved and persisted in the fields of philosophy and science for a long time, one of the greatest of its challenges is to provide a proper definition of this concept. The term 'consciousness' is often used in literature with different meanings, which often impedes communication between and even within different areas of study (Rosenthal, 2009). The next section will give an overview of the most important perspectives on what consciousness is, starting from the philosophical tradition and ending with clinical aspects of relevance to this thesis.

1.3.1. QUALIA

One term intimately related to consciousness is *qualia*. It refers to subjective experiences such as seeing red, smelling a rose or tasting chocolate. It does not refer to a technical account of how these experiences arise with the mediation of sensory pathways, but to the phenomenal character of things in the world as we experience them introspectively (Jackson, 1982).

A thought experiment that can clarify this concept is given by the argument of inverted qualia (John Locke, 1689). Assume that when one person sees the colour red, another person sees what the former person would classify as the colour blue (and vice versa). Both persons will identify the colour as 'red' and the neural pathways starting from the retina, passing through the visual cortex and ending in the system that controls the reporting of the colour red will be precisely the same in both. Still, the experience accompanying the same pattern of neuron firing will be different in the two persons. The inverted qualia will also not be detectable, since it is not possible to directly convey the experience of a person to another, and the reports they produce will match.

Daniel Dennett summarises four essential properties of qualia (Dennett, 1988): they are ineffable, intrinsic, private, and directly or immediately apprehensible in consciousness. In the same work, Dennett dismisses qualia as introspective illusions. He is not the only philosopher who does so (James, 1904). Still, many others endorse the existence of qualia as valid phenomena (Block, 1995; Chalmers, 1995; Jackson, 1982; Searle, 1997). The latter view seems to be prevalent in the neuroscientific community, although, as will be explained below, there are certain limits to what the scientific method can do with regards to this concept (Tsuchiya, 2017).

1.3.2. WHAT-IT-IS-LIKENESS

One of the most widely-used definitions of consciousness comes from Thomas Nagel, who wondered 'what is it like to be a bat' (Nagel, 1974). There are several difficulties we encounter if we try to grasp the subjective experience of being a bat. Although we might be able to contemplate the notion of

having webbed wings and perceiving the world through echolocation, we do so by using our imagination, which is rooted in data we have from our own human senses only; we hence can only imagine what it would be like for a *human* to behave like a bat, not for a *bat* to be a bat. This highlights a key aspect of conscious experience: its *first-person, subjective, private* character.

The idea of what-it-is-likeness has often been used by prominent neuroscientists and philosophers to define consciousness (Block, 1995; Chalmers, 1996; Tononi, 2008; Tye, 1992). Although intuitively useful, this definition is currently problematic in terms of providing a rigorous scientific explanandum. Empirical science would instead require a *third-person object* susceptible to some kind of measurement, while what-it-is-likeness describes a reflexive relationship that we can only know introspectively. One way to better inspect the semantic meaning of the expression 'what it is like' is to translate it into other languages. One observes that in multiple languages it translates to a 'how'-question (Stoljar, 2016), which suggests a reference to a property of an object, as opposed to an object itself. In this sense, what-it-is-likeness is reflexive or self-referential because it seems to describe consciousness as a property of consciousness. Furthermore, what-it-is-likeness might require a 'homunculus' that experiences the state of 'what it is like' to be something, potentially causing a recursive self-referential problem. Therefore, Nagel's definition is intuitively useful, but less so scientifically, as it does not provide us with a measurable entity.

1.3.3. MARY THE NEUROSCIENTIST

In the same publication, Nagel also argues that mental properties cannot be reduced to physical properties, or at least we do not yet have an understanding of how this could happen. Although the aim of this section is not to provide arguments for or against certain philosophical positions on consciousness, some of such arguments are useful when trying to pin down what the concept of consciousness refers to and what our questions are. Another persuasive argument against physicalism comes from Frank Jackson, who introduces a fictional neuroscientist called Mary (Jackson, 1982).

Mary studies human vision, for whatever reason, through a black-and-white television screen, from a black-and-white room. In other words, she has never seen colour. Nonetheless, being a brilliant scientist, she has managed to learn everything there is to know about colour. For example, she can describe in detail the exact properties of light that need to be fulfilled and the precise pathway through the retina, visual cortex and so on, leading to a person verbally reporting that they see the colour red. Jackson asks what happens when Mary gets out of the room and sees the colour red in reality for the first time. Does she learn something new? Jackson believes that she does, which justifies the existence of a mental property that is not explicable by a physical account only. Still, others have argued she

does not (Dennett, 2007) or that what she learns can be accounted for by a physical explanation (Lewis, 1999).

Regardless of one's position in this debate, this thought experiment brings up once again two fundamental topics of consciousness: on the one hand, the quest for understanding the physical mechanisms that underlie consciousness from an *objective, third-person* point of view; on the other hand, the *subjective, first-person* mental experience.

1.3.4. THE EXPLANATORY GAP

In the scientific quest to understand consciousness, the difference between the subjective and the objective aspects of consciousness is easily overlooked. For example, some papers start by invoking the ineffable aspect of consciousness and continue by proposing a solution to it, but this solution refers in fact to a related but different problem, such as that of reportability or introspective access (Chalmers, 1995). This is understandable: scientific progress requires measurable properties (Maxwell, 2004) – even when this only consists of subjects reporting their experiences. In the case of consciousness as we currently conceive of it, scientific approach has no choice but to avoid the unmeasurable aspect. Conscious experience must be equated with behaviour or with the content of linguistic reports. It is indeed imperative to continue using this approach in order to keep making progress of practical importance. Medicine is one major field that greatly benefits from this approach, as will be emphasised in more detail throughout this thesis.

Still, if only for the sake of epistemological honesty, it must always be acknowledged that the scientific method does leave unsolved the mystery of conscious experience, as it has propagated through ages, from Aristotle through to Descartes. Joseph Levine, despite expressively declaring his materialist stance, argues for what he calls the 'explanatory gap' (Levine, 1983). In his work, he examines the contingency of the statements 'pain is the firing of C-fibers' and 'heat is the motion of molecules'. It is conceivable that both propositions could, under certain conditions, be false. However, upon further inspection, the contingency of the proposition about heat can be explained away by providing a proper definition for heat. On the other hand, in the proposition about pain, an explanatory gap persists due to the reference to a *qualitative*, *subjective* property. In his work, Levine concludes by arguing that the only way to eliminate the mind-body problem, which he deems unsolvable within the framework of materialism, is to simply reject our intuitions about the existence of qualia.

1.3.5. THE HARD PROBLEM

Not all philosophers embrace an ideology as gloomy as Levine. David Chalmers (Chalmers, 1995, 1996) designates the problem on the subjective side of Levine's explanatory gap as the *hard problem* of

consciousness: the question of why the firing of cells in our neural systems, for example the firing of C-fibres, should be accompanied by any experience, in this case that of pain. This holds for any other modality, such as vision: if we could dissect a brain reporting that it is experiencing the colour blue, we would only find neural machinery, but not the colour blue. Where is the experience of colour then, and how does it arise from our colourless brain?

Ned Block uses the term *phenomenal consciousness* (P-consciousness) to refer to experience (Block, 1995). His definition partially overlaps with the concept of experience as used by Chalmers. However, at the same time, Block assumes that P-consciousness is just another concept of consciousness implemented by the brain. The implementation of P-consciousness in the brain might or might not overlap that of *access-consciousness* (A-consciousness), which covers the aspects of consciousness which are reportable. On the other hand, the hard problem of consciousness refers strictly to the aspect of consciousness which is not captured by any functional description of the neural system.

This problem is 'hard' because we do not currently know how to approach it and what an answer might look like. Some, such as Colin McGinn, argue that it will always remain a mystery, by virtue of our very cognitive structure and limitations (McGinn, 1989). Others, however, consider that in absence of a direct way to tackle the hard problem, we should for now concentrate on the so-called easy problem of consciousness, presented below. Eventually, progress on it might shed some light on the hard problem (Crick and Koch, 1998).

1.3.6. THE EASY PROBLEM

The *easy problem* of consciousness is not easy, but at least we know how to approach it using the empirical method. The ultimate goal is the provision of a full account of how the human neural machinery supports and gives rise to any processes we classify as conscious (Block, 1996; Chalmers, 2000; Rees et al., 2002). The hard problem implies that there is a leap of faith that needs to be taken to link true experience – the topic of the hard problem – to its report. There is no direct evidence of the fact that what seems to us to be awareness in another human being based on behaviour or reports is indeed true experience; however, it is sensible to assume so, and this assumption will be made henceforth in most of this thesis.

Neuroimaging tools, such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), can be used to identify neural correlates of consciousness. By experimentally varying an element of consciousness, we can compare and contrast brain activity in different conditions, thereby revealing specific mechanisms that are required for consciousness. Moreover, using neurostimulation techniques such as transcranial magnetic stimulation (TMS) to briefly induce a

change in brain activity, we can investigate the causal relationship between neurodynamics and consciousness. Chapter 2 is devoted to a review of how such techniques have been advancing our knowledge related to consciousness and have subserved valuable clinical applications.

1.3.7. DIMENSIONS OF CONSCIOUSNESS

In investigating the relation between brain activity and consciousness as described above, researchers commonly distinguish between two complementary dimensions that are especially relevant in the clinical assessment of disorders of consciousness: level (or state) and content (Laureys, 2005; Rees et al., 2002). The level of consciousness refers to the degree of wakefulness, whereas the content refers to the extent that someone is aware of specific stimuli. The level of consciousness represents an intransitive formulation: being conscious, as opposed to being asleep, anaesthetised, or comatose. The content of consciousness refers to a transitive usage: being conscious *of* something, either of a perception or of a higher-order thought. Qualia refers most commonly to the content of consciousness (Tye, 1992).

The content and the level of consciousness are conceptually distinct, but not fully orthogonal dimensions (Hohwy, 2009; Overgaard and Overgaard, 2010). A high level of consciousness, i.e. being awake, is normally required in order to entertain any content of consciousness. Dreaming, however, – rich content in the absence of wakefulness (Siclari et al., 2013) – occurs in the absence of wakefulness. Conversely, being awake usually implies some content of consciousness, even during boredom (Eastwood et al., 2012), but patients with disorders of consciousness seem to have sleep-wake cycles without showing any signs of awareness (Bekinschtein et al., 2009).

1.3.8. FROM PHENOMENOLOGY TO BEHAVIOUR

Can we bridge the gap between the hard and the easy problems of consciousness in the study of either of these dimensions? It has been argued (Seth, 2016) that we should adopt a balanced perspective and attempt to solve what is currently the real problem of consciousness: accounting for the phenomenological properties of consciousness, as discovered through both introspection and objective measurement, by mapping them onto biological mechanisms. This perspective allows us to construct a comprehensive picture of first-person consciousness, while leaving aside the worry regarding why consciousness exists in the first place. Such an empirical approach can thus give us a key to solve the introspective mysteries of conscious experience. In following this approach to probe the content of consciousness, we may correlate neural activity with verbal or behavioural reports of experiences such as whether a stimulus was seen or not, or the orientation of a Gabor patch briefly

flashing on the screen. However, some concerns related to these approaches exist and solutions to them have gradually began to be proposed.

First, laboratory experiments usually employ minimalistic stimuli, which do not approximate well the richness of conscious experience (Naci et al., 2014). In an attempt to solve this problem, a renewed focus on incorporating full phenomenological reports into empirical science has been suggested (Zahavi, 2004). For example, neurophenomenology (Varela, 1996) proposes that a systematic exploration of the structure of human experience using both subjective reports and objective evidence can provide a remedy for the hard problem. Other frameworks have also been proposed; for example, under the view of predictive coding (Clark, 2013; Friston and Kiebel, 2009), the brain continuously predicts the world, while sensory inputs only provide the data for correcting the predictions and adjusting priors about the world. Under this view, conscious experience is an ongoing, controlled hallucination, with minimal stimuli required to study how perception is affected by disrupting predictions at different levels in the cortical hierarchy (Seth, 2016).

Secondly, when a subject is asked to make a report, he has to perform an act of introspection that also requires other cognitive functions, such as attention (Lamme, 2003), or working memory (Soto and Silvanto, 2014). It is therefore challenging to subsequently disentangle the neural correlate of the reported phenomenon and that of the cognitive acts required for reporting. To address this, a recent key development has been the introduction of no-report paradigms (Tsuchiya et al., 2015). These paradigms are designed to allow bypassing reports when inferring whether a subject is perceiving a certain stimulus. For example, in binocular rivalry, perceptual switches can be inferred from eye movements (Frassle et al., 2014). When carefully used, such developments of experimental design allow researchers to get closer to what subjects are truly aware of at a given moment.

These approaches are useful particularly in assessing the content of consciousness in the presence of wakefulness. In contrast, to investigate the level of consciousness, it is more difficult to obtain subjective reports, especially in pathological cases like disorders of consciousness. Furthermore, in cases of natural loss of consciousness, such as the onset of sleep, providing the reports interferes with the process of losing wakefulness. To circumvent this problem, *passive* methods of approximating wakefulness are sometimes used. For example, the onset of sleep is accompanied by a variety of physiological markers: changes in electrical brain activity and connectivity measured at rest (Scammell et al., 2017) or in response to external stimuli (Chennu and Bekinschtein, 2012), but also body measures such as muscle relaxation or the slowing down of breathing (Goupil and Bekinschtein, 2012; Ogilvie, 2001). On the other hand, for healthy subjects, careful *active* methods can also be used with minimal interference into the transition into unconsciousness. For instance, subjects may be

repeatedly woken up and asked whether they had been awake immediately prior to the awakening (Hori et al., 1994), or they might perform a task requiring simple button presses during anaesthetic induction or in a setting that encourages them to fall asleep (Kouider et al., 2014). Such simple reports provide a minimal window into first-person experience, but they allow us to understand the neurodynamics underlying a dramatic aspect of our mental world and a fundamental behaviour: the loss and recovery of consciousness.

1.4. TRANSITIONS OF CONSCIOUSNESS

How can it be that from one moment to another, as we fall into deep sleep or ingest an anaesthetic drug, we temporarily drift into the realm of unconsciousness, where are no longer able to access our vivid mental world? Unconsciousness is a fascinating territory that, unlike consciousness, cannot be scrutinised through introspection. As consciousness is interrupted, brain activity is still ongoing, but it produces distinct signatures compared to wakefulness (Schwartz et al., 2010). Some neural markers are similar across different conditions of unconsciousness, like sleep and anaesthesia (Murphy et al., 2011). Can we pinpoint the neural boundary where consciousness disappears by studying the transitions between levels of consciousness?

The idea of levels of consciousness does not necessarily imply a theoretical ordering of various states of consciousness along an axis representing the 'intensity' of consciousness (Bayne et al., 2016). Historically, the level of consciousness has been introduced in conjunction with disorders of consciousness (Laureys, 2005) to emphasise a gradation between comatose, vegetative, minimally conscious and healthy subjects. In that context, the term referred to the trajectory of clinical recovery starting from the most severe form of impaired consciousness and ending with full recovery (Laureys and Tononi, 2011). However, within each of these pathological states, a direct comparison between individuals is not theoretically sound, as individual pathologies lead to different palettes of preserved cognitive functions (Fernández-Espejo and Owen, 2013). Similarly, there is no ranking implied by the concept of consciousness levels between sleep, anaesthesia or other forms of altered consciousness. By contrast, each of these states defines a unique *transition* between consciousness and unconsciousness, thereby providing a framework for determining what is lost and regained between these two states.

This thesis explores the delicate boundary between consciousness and unconsciousness by contrasting wakefulness with three conditions of impaired consciousness: sleep, sedation and coma. In the following section, a brief introduction to each of these states of consciousness will be made. It will be argued that advancing our understanding of these fields is not only essential in clarifying the

mystery of consciousness, but also in the advancement of medical tools that can assess and monitor the evolution of a wide range of clinical conditions, such as diagnosing and treating insomnia and disorders of consciousness, or optimally maintaining anaesthesia during surgery. Ultimately, the mystery of first-person consciousness might, by definition, not be solvable using the empirical method; however, elucidating the neural substrate underlying different levels of consciousness is a *real*, solvable problem that will bring immediate benefits to society.

1.4.1. WAKEFULNESS AND SLEEP

Along with the rising and setting of the sun, sleep fragments the introspective film of our lives into days. Every day, for around seven hours – if appropriate guidelines are followed (Watson et al., 2015) – we lose the rich experience of our surroundings, knowing with certainty that we will, with no effort, regain it in the morning. Sleep is, however, not fully devoid of experience. As we enter sleep or wake up, we may perceive glimpses of mesmerising or puzzling hypnagogic imagery or intrusions (Noreika et al., 2015). As we go into more profound sleep, we dream (Siclari et al., 2013). What are the neural dynamics that underlie the familiar descent into the alien territory of unconsciousness?

Sleep provides a hallmark reference for the definition of consciousness. As science still struggles for a rigorous definition, it is common to refer to consciousness as the familiar, sentient state in-between awaking from dreamless sleep and falling asleep (or becoming otherwise 'unconscious') (Searle, 1993). In a sense, the loss of consciousness that sleep provides on a daily basis is a privilege: if we were fully conscious in every moment of our lives, we might take consciousness for granted and assume it is a natural by-product of any healthy human brain. Sleep proves to us that this is not the case: a fully-functioning human brain can indeed be unconscious.

Fundamentally, sleep is a *reversible* behaviour marked by quiescence and an elevated arousal threshold (Vassalli and Dijk, 2009) only found in higher evolved organisms (Hobson, 1995). Sleep is necessary: deprivation causes a range of metabolic (Knutson et al., 2007) and cognitive (Harrison and Horne, 2000; Pilcher and Huffcutt, 1996; Tsai et al., 2005) impairments, as well as an increased need for sleep (Berger and Oswald, 1962). While recent literature has been exploring the functions of (access-) consciousness (Samaha, 2015), sleep shows that unconsciousness is associated with essential functions in humans too, such as learning and memory consolidation (Hobson and Pace-Schott, 2002; Stickgold and Walker, 2007), potentially by downscaling the synaptic gain that occurs during wakefulness (de Vivo et al., 2017; Tononi and Cirelli, 2006).

Sleep occurs as an alteration, rather than a cessation, of brain activity (Tononi and Massimini, 2008). The EEG has been extensively used to monitor neural activity during sleep for nearly a century

(Dement and Kleitman, 1957; Loomis et al., 1935) and a general architecture of human sleep patterns has emerged. Currently, the most widely-used scheme is provided by the American Academy of Sleep Medicine (AASM) (Iber et al., 2007), according to which the EEG recording is divided into 30-second epochs that are then manually (Silber et al., 2007) or automatically (Ronzhina et al., 2012) classified into five categories: wakefulness, NREM (N) stages 1 to 3, and REM sleep. Specific EEG markers include the presence of more than 50% alpha (8-13 Hz) waves for wakefulness, theta (4-7 Hz) and sharp vertex waves for N1, k-complexes and spindles for N2, slow waves (0.5-2 Hz) for N3, and rapid eye movements with mixed wave activity for REM (these EEG patterns are described in more detail in Chapter 2). But how can we capture the moment when we actually lose consciousness? Most researchers establish that the true onset of sleep occurs when N1 ends and N2 starts (Ogilvie, 2001), although some studies report that less than half of subjects perceive themselves as being asleep at the beginning of stage 2 (Hori et al., 1994; Sewitch, 1984). To explore the electrical dynamics of the process of falling asleep, the Hori scoring scheme provides nine finer-grained levels of scoring, with the last level corresponding to the beginning of N2, applicable to periods of a few seconds (Hori et al., 1994).

But are even a few seconds enough to capture the intricate subjective and objective changes in consciousness as we fall asleep? It has been shown that the electrical field of the brain exhibits periods of quasi-stability, that last as little as tens of milliseconds (Khanna et al., 2015; Koenig et al., 2002), whose dynamics are altered during sleep (Brodbeck et al., 2012). Networks of the brain fluctuate at millisecond level during a resting state (Baker et al., 2014). This suggests that understanding the millisecond-level fluctuations in neural activity may give us an even better view of how consciousness is lost at the onset of sleep, as will be further explored in Chapter 3.

In addition to the EEG markers of sleep stages, there are other typical changes that happen as we lose consciousness. For example, in spectral domain, alpha power and connectivity disappear, as lower and higher power and connectivity emerge (De Gennaro et al., 2016; Hudetz et al., 2015). Similar changes are also observed in the pathological loss of consciousness (Chennu et al., 2017, 2016a, 2014). But what are the dynamics of fast-changing microstates and brain networks as we fall asleep? Chapter 3 will combine the method of microstates and that of spectral power and connectivity to reveal the swift changes occurring in the EEG during the process of falling asleep, and suggest that fine-grained dynamics provide a valuable window into the transition to unconsciousness.

1.4.2. SEDATION AND ANAESTHESIA

The idea of unconsciousness can be frightening. Still, at rare occasions, we might want to be temporarily unconscious. For example, we want to avoid the intense pain of surgical procedures.

Humankind has known about natural hallucinogenic agents that alter the state of consciousness for thousands of years (Garcia-Romeu et al., 2016), but it was only in the 19th century that drugs producing temporary unconsciousness were discovered (Bigelow, 1846). Since then, drugs like propofol, ketamine or halothane have provided a reliable clinical solution for general and local anaesthesia (Brown et al., 2011), with undesired awareness during anaesthesia estimated to occur in as few as 0.13% of the cases (Sebel et al., 2004). In the case of milder procedures such as dental extractions, a state of sedation, where the central nervous system is depressed but the patient is still responsive, can be preferred (Lyratzopoulos (Liratsopulos) and Blain, 2003).

How do these drugs act on the nervous system to extinguish consciousness? A common mechanism for anaesthesia is still under debate (Alkire et al., 2008; Mashour, 2004). To an extent, anaesthesia and sleep display common neural signatures (Franks and Zecharia, 2011; Murphy et al., 2011; Schwartz et al., 2010). At molecular level, different anaesthetics have a variety of effects. For example, propofol binds to the $GABA_A$ receptors, thereby acting as an inhibitor on the nervous system, including on the connections between the thalamus and the cortex. This might be the mechanism that causes unconsciousness (Bai et al., 1999; Brown et al., 2011). Other drugs affect different receptors, such as NMDA (ketamine), dopamine (droperidol, morphine), or opioids (morphine). Furthermore, different drugs affect various brain regions differently. For example, at similar level of behavioural effect, propofol decreases blood flow in the frontal brain areas, whereas thiopental decreases blood flow in posterior and cerebellar areas (Veselis et al., 2004). With a variety of local changes that are associated with unconsciousness for each drug, a better explanation might arise from the higher-level dynamics of neural activity.

One potential answer lies in the observation that all anaesthetics seem to disrupt global integration of neural activity (Alkire et al., 2008). To some extent, anaesthetics also disrupt local neural activity differentiation (Schartner et al., 2015). Intriguingly, at sedative doses of anaesthetic, the point where responsiveness is lost does not only depend on the level of drug, but also on other individual factors, such as the strength of global alpha brain networks (Chennu et al., 2016a). This agrees with several theories of consciousness, as will be detailed in Section 1.6. In particular, information-theoretic measures of integration (King et al., 2013) and differentiation (Schartner et al., 2015) in the EEG might be promising candidates for better understanding the global neural effects that explain the loss of consciousness. Therefore, Chapter 4 of this thesis will investigate information-theoretical measures of integration and differentiation during moderate sedation with propofol, at the point where responsiveness is lost in some (but not all) patients. It will be shown that responsiveness and drug exposure have overlapping but distinct neural signatures, which suggests that the loss of

consciousness might indeed be explained by the dynamics occurring at high-level organisation of the neural system.

1.4.3. DISORDERS OF CONSCIOUSNESS

Severe brain injury can rob us of consciousness, temporarily or forever. All that is required for this grim prospect is damaging any part of the evolutionarily older neural infrastructure that supports basic biological and cognitive functions, as detailed in Section 1.2. This can be caused by trauma to the head or by non-traumatic causes, such as haemorrhage or ischaemia (Bagnato et al., 2010).

Patients surviving brain injury typically go through a sequence of progressive stages towards recovery (Laureys et al., 2005). *Coma* is closest to brain death: the absence of awareness and wakefulness. Patients in a coma typically only exhibit reflex activities mediated by the brainstem, but not by the cortex, and they cannot be aroused even by strong and obnoxious stimuli (Laureys and Tononi, 2011). A sign of recovery from coma and entering a *vegetative state* is the reappearance of wake and sleep cycles. At this stage, the patient can be aroused, despite no signs of awareness of the external world. If the vegetative state persists for more than a month, recovery becomes unlikely; for example, only around 20% of persistent anoxic vegetative patients will regain responsiveness within two years (Estraneo et al., 2013). Finally, a *minimally conscious state* is diagnosed when there are limited or inconsistent signs of awareness, from which the patient might progress to a recovery of consciousness.

Clinical diagnosis in disorders of consciousness is not straightforward. Assessment methods like the Glasgow Coma Scale (Jones, 1979), the Coma Recovery Scale (CRS) (Giacino et al., 1991), or the Coma Recovery Scale Revised (CRS-R) (Giacino et al., 2004) measure *behavioural* responsiveness to account for the degree of processing of different types, such as visual, auditory, motor, communication or arousal. A score is given on each of these subscales, reflecting how elaborate the responses of the patient are, and the final score places the patient on a scale used for clinical evaluation in conjunction with other biological measurements. However, based on such scales, more than 40% of vegetative state patients have reportedly been misdiagnosed, as they show signs of minimal consciousness when evaluated by expert teams (Schnakers et al., 2009). In acute coma, patient outcome at individual level can still not be accurately predicted (Stevens and Sutter, 2013).

Further underlining the need to improve behavioural assessment in these patients, it was recently discovered that patients classified as vegetative may in fact have preserved, but covert, cognitive functions. They might activate particular brain areas in response to command, despite their inability to make overt responses (Fernández-Espejo and Owen, 2013). In a pioneering study by Adrian Owen and colleagues, it was reported that a vegetative patient scanned using fMRI produced brain activity

in the same areas as healthy subjects when asked to play tennis and to imagine walking around her house (Owen et al., 2006). This finding has been replicated in a number of vegetative patients using fMRI (Monti et al., 2010) and bedside EEG (Cruse et al., 2011; Gibson et al., 2014). A most likely explanation would be that these patients are at least partially conscious, but unable to make overt responses. An alternative explanation would be that these are complex but automatic responses, that demonstrate a preservation of cognitive functions, which can be, however, independent of conscious experience (Overgaard and Overgaard, 2011). Indeed, complex responses, such as conflict resolution and response inhibition, have been shown to occur in the absence of first-person awareness (Van Gaal and Lamme, 2012). Nevertheless, the discovery that vegetative patients are able to activate brain areas similarly to healthy subjects in response to command shows that such patients have in place at least the potential architecture for supporting a subjective inner experience, in absence of the ability to respond.

To address this, a number of electrophysiological signatures have recently been proposed for aiding clinical diagnosis in disorders of consciousness (Sitt et al., 2014). One promising such tool is probing the strength of spectral brain networks, which have been shown to undergo characteristic changes in both disorders of consciousness (Chennu et al., 2017, 2014) and sedation (Chennu et al., 2016a). By examining the topology of EEG brain networks at frequencies of interest and evaluating the efficiency of local and long-range neural activity and connectivity, an informative index could be obtained in order to assess the clinical state of a patient. Can this approach be useful for prognostication in the acute phase of coma? In Chapter 5, graph-theoretical measures will be employed on the EEG networks of comatose patients with recent traumatic brain injury in an attempt to find key signatures that might predict their eventual outcome. Although this approach provides only limited insight into the philosophical concept of first-person consciousness, it demonstrates how the neuroscientific study of particular elements of consciousness can provide a clinical methodology that can make a life-saving difference to patients fighting to regain consciousness in the intensive care unit.

1.5. RESPONSIVENESS AS A PROXY FOR CONSCIOUSNESS

In the search for neural signatures that indicate a state of consciousness or its absence, this thesis uses measures of behavioural responsiveness in order to zoom into the space that encompasses the transition between states of consciousness. Chapter 3 employs the responsiveness to a semantic categorisation task as a covariate of the conscious state during the process of falling asleep. Chapter 4 uses the responsiveness to a perceptual discrimination task during propofol-induced unresponsiveness. In both cases, as will be shown later, there is a gradual increase, as the experiment

progresses, in the number of misses and sometimes in reaction times, validating a pattern of transition towards unconsciousness. Finally, Chapter 5 employs the CRS-R behavioural score to measure the degree of recovery in comatose patients two months after traumatic brain injury. In all these experiments, behaviour is a *proxy* for measuring consciousness.

This proxy is, however, not perfect. Indeed, responsiveness does not imply consciousness. First, in healthy adults, classical conditioning is a simple example of responsiveness that does not require awareness (Clark and Squire, 1998). Secondly, some patients with lesions in their primary visual cortex exhibit blindsight, where they are able to respond to stimuli they deny being aware of (Zucco et al., 2014). Finally, it has been argued that split brain patients respond to stimuli presented to their right hemisphere, and later may create an alternative explanation that justifies the response, demonstrating that the stimulus was not consciously perceived (Gazzaniga et al., 2014). This example should, however, be interpreted with caution, due to the isolated character of such case studies. More recently, it has been argued that such results could be better explained by a unified consciousness experiencing two unintegrated perceptual streams that are difficult to integrate (Pinto et al., 2017).

Conversely, unresponsiveness does not imply unconsciousness. First, some subjects undergoing general anaesthesia are able to communicate using one forearm isolated from the anaesthetic, despite not having ulterior recollection of this (Sanders et al., 2012). Secondly, patients classified as vegetative might produce neural responses similar to healthy adults, as discussed in the previous section (Cruse et al., 2011; Monti et al., 2010). This allows the possibility that consciousness can occur in disorders of consciousness without any overt responsiveness (Alkire et al., 2008; Boly et al., 2013a; Sanders et al., 2012).

Keeping this in mind, responsiveness uncontroversially provides a window into a *segment* of the whole space between consciousness and unconsciousness. In the case of falling asleep and entering sedation, responses are objective, momentary measures of the ability of the neural system to fully engage with a stimulus and produce a response. Introspectively, the loss of consciousness does not seem to be a binary event, but rather a gradual process with intermediate states between full wakefulness and sleep (Sewitch, 1984). The responsiveness approach probes the particular part of this intermediate state where responsiveness is lost. Moreover, this may have valuable applications, such as preventing driving accidents related to falling asleep at the wheel (Horne and Reyner, 1999) or improved individual dosage in anaesthesia (Alkire et al., 2008). In the case of coma, the CRS-R score provides an estimated level on the behavioural scale that a patient needs to climb in order to achieve the cognitive functions that support healthy consciousness. In this latter case, subjective consciousness matters less, while the clinical signs of improvement are essential in assessing recovery.

Hence, while not perfect, responsiveness provides a unique and valuable marker in the investigation of healthy and clinical transitions between levels of consciousness.

1.6. THEORIES OF CONSCIOUSNESS

What could a theory of consciousness look like? A practical alternative to the dualist proposition that consciousness is ontologically separate from physical matter (Chalmers, 1995) or to deeming the problem unsolvable (McGinn, 1989) are identity theories that equate neural events with phenomenology. While an explanatory gap remains in this case, such theories provide a way to approach this topic and enhance our understanding of mental and neural processes related to consciousness. Moreover, in clinical settings, these approaches can also provide helpful frameworks to assess patients with disorders of consciousness or to monitor healthy subjects undergoing anaesthesia. Importantly, most modern such theories have stepped beyond being *biological theories* (Block, 2009), which straightforwardly claim that activity in certain fixed brain areas might give rise to consciousness. Instead, they propose specific *global patterns* of neural activity that are responsible for conscious experience. Three theories of this type that have led to valuable advances will be presented below. The purpose of this section is not to provide a comprehensive review of theories of consciousness, but to focus on theories of interest for the further chapters of this thesis.

1.6.1. GLOBAL WORKSPACE THEORY

The global workspace theory (Baars, 1988) builds upon the observation that consciousness encompasses a momentarily unified collection of information processed by specialised mental modules. The theory hence frames consciousness as a broadcasting signal that is globally accessible in the brain. As highly-specialised modules perform different computations, which are intrinsically unconscious, they may momentarily become part of the *global workspace* and thereby share information with other submodules of the brain. When this happens, the theory posits that we are conscious of the globally shared content. A key remark is that there is a considerable amount of neural processing occurring at any given time in the brain, which we are not aware of. This theory has been further extended into the global neuronal workspace theory (Dehaene and Naccache, 2001), which emphasises the role of top-down attention in mobilising the relevant neuronal modules that become available in the global workspace. Key brain regions whose global implication in neural activity is consistent with this theory include frontoparietal and medial temporal areas (Baars, 2005). The global workspace framework has made useful predictions that can be used in both clinical and research contexts, such as the relevance of long-distance information sharing in consciously processed information (Dehaene and Changeux, 2011). One disadvantage of this theory is that it does not offer

an explanation of the phenomenological structure of subjective experience, but simply suggests how information sharing in an interconnected system might be implemented.

1.6.2. DYNAMIC CORE HYPOTHESIS

The dynamic core hypothesis (Tononi and Edelman, 1998) makes a further conceptual step by identifying two fundamental aspects of conscious experience: *integration* and *differentiation*. Integration signifies that every single scene of our stream of consciousness is perceived as a whole and is not separable into a sum of components or into multiple points of view. Differentiation refers to the complexity of every scene we experience, which is unique among an unimaginably vast number of other possible experiences. These two phenomenal properties are hypothesised to map to integration and differentiation of measured brain activity arising from different neural modules with specialised functions. In particular, re-entrant thalamo-cortical connections might be essential for integration (Seth and Baars, 2005), creating the concept of consciousness as a 'remembered present' (Edelman, 2001). The theory underlines that there is no need for a specific subset of brain areas where integration and differentiation are implemented, but suggests a 'dynamic core' of neural modules that may vary across time and across people, which, in the right configuration, gives rise to consciousness.

By suggesting two specific properties of phenomenal experience that may directly map to characteristics of brain activity, this approach is useful for designing neurophysiological measures that can be applied in order to track the state of consciousness of patients under anaesthesia or with disorders of consciousness. Several such measures will be presented in Chapter 2.

1.6.3. CONSCIOUSNESS AS INTEGRATED INFORMATION

In contrast with other neuroscientific theories that start from neural events in an attempt to describe their effect on conscious experience, the integrated information theory (IIT) (Oizumi et al., 2014; Tononi, 2004; Tononi et al., 2016), attempts to build an explanatory bridge between subjective experience and brain activity by starting from phenomenology itself. It first proposes a set of axioms – self-evident truths that describe conscious experience. The most recent version of IIT (Oizumi et al., 2014) proposes that first-person perspective is described axiomatically by the facts that consciousness *exists*, that it consists of a *composition* of multiple elements that are experienced at the same time, that it is *informative* due to its distinctiveness from any other possible experience, is *integrated* and irreducible to the sum of its components, and *exclusive* of other simultaneous conscious experiences. The theory then proposes a set of postulates that specify the laws that a physical system must satisfy to give rise to consciousness as described by these axioms. The postulates parallel the axioms. As a starting point, the existence axiom is translated into the postulate that a mechanism can contribute

to consciousness only if it generates *cause and effect* within a system. Further, these mechanisms can be *composed* and *integrated* in a way that is not reducible to the repertoire of cause and effect of its independent components. Moreover, the theory proposes that consciousness excludes other experiences by emerging only at the mechanistic level where the integration between components, ϕ, is maximal.

Several definitions of ϕ exist, which vary with specific assumptions made regarding the system for which integrated information is computed. For example, early measures were applicable only discrete Markovian dynamics, but more recent measures are applicable to any stochastic continuous dynamics (Barrett and Seth, 2011; Kim et al., 2018; Mediano et al., 2018). However, these are challenging to compute due to the explosive computational complexity required for a system proportional with the neural system. Instead, measures that approximate the constraints required by the postulates can be used. One of the predictions of the theory is that sleep, anaesthesia and disorders of consciousness are characterised by low or absent consciousness due to the loss of neural information integration and differentiation. Measures inspired by information integration theory, which will be described in the next chapter, have been found to comply with this prediction (Casali et al., 2013; Massimini et al., 2005; Seth et al., 2008). Conceptually, the theory predicts that feedforward systems that achieve the same functionality as a complex integrated system, but do not generate cause and effect as a whole, are not conscious. On the other hand, the mathematics of the theory, which do not distinguish between biological and other systems, mean that IIT, by itself, is compatible with a philosophical framework of panpsychism (Tononi and Koch, 2015). Any system composed of interacting units that generates a degree of cause-and-effect properties can be described using the measure of consciousness ϕ. Would that mean that some of our computers are conscious? Creative speculation may be entertaining, but we do not know how such predictions could be testable.

Nevertheless, the information integration theory currently offers perhaps the most promising perspective of bridging the gap between phenomenology and neuroscience (Tsuchiya, 2017). This theory is still in its early days and more work is needed to further develop it into applicable forms for providing a solid foundation for the progress of consciousness research.

1.7. NEXT CHAPTERS

In the quest of investigating the boundaries of consciousness levels, this thesis will experimentally explore the three transitions presented in section 1.4: sleep, sedation, and coma. Considering the specificities of each of these transitions and the developments in pre-existing literature, a distinct question will be asked in each case and different analysis methods will be applied, as appropriate.

First, for the sleep analysis in Chapter 3. , the investigation concerns the fast-paced dynamics of the transition to unconsciousness. One motivation for this analysis is the need for finer-grained measures for assessing the transition to sleep. As will be detailed below, standard sleep scoring systems use 30-second epochs to classify EEG data during the process of falling asleep (Iber et al., 2007). However, the transition to unconsciousness can be further broken down into the 4-second Hori system, which is consistent across subjects, but is still not fully congruent with respect to subjective reports (Hori et al., 1994). The key to understanding the transition to sleep might lie in the millisecond-level dynamics of the neural system. Therefore, this chapter focuses on the rapid changes present in the electroencephalogram before and after the loss of responsiveness during drowsiness, as captured by electric microstates of the brain, and complemented by an analysis of spectral power and connectivity changes at the same rapid temporal scale.

Secondly, in Chapter 4. , the dynamics of brain activity during sedation are analysed using measures inspired by information integration theory. The dataset used in this chapter was previously analysed using spectral measures (Chennu et al., 2016a), hence the information-theoretical analysis aims to bring a complementary perspective at the same temporal scale as the previous analysis. The Lempel-Ziv complexity (Lempel and Ziv, 1976) of the electroencephalogram is measured to quantify local dynamics, while the weighted symbolic mutual information index (King et al., 2013) is used to assess connectivity. An analysis of electric microstates would also be interesting to apply to this dataset; however, at the time of this study, we were aware of studies already in progress directly investigating the same question (Britz, 2015), so we decided to focus on a novel analysis inspired by the theory of information integration.

Finally, the analysis of comatose patients in Chapter 5. focuses on the restoration of brain networks in the acute phase of traumatic brain injury. This analysis is more clinically-oriented compared to the sleep and sedation studies. In comatose patients, the disruption of brain activity is reflected in a slowing-down of informative EEG frequencies, hence an analysis of delta, theta and alpha frequency networks is performed. Moreover, due to the alteration of brain geometry, a topographical analysis as performed by the method of electric microstates would not be appropriate in these patients. Markers of spectral power have been extensively already described in coma literature (Kaplan, 2004; Lehembre et al., 2012a, 2012b; Thatcher et al., 1991). This chapter proposes a new methodological approach – that of graph theory applied to connectivity networks – that could to aid diagnosis and prognosis after traumatic brain injury.

By applying a specific set of analyses as appropriate for the question explored within each of these transitions, this thesis also demonstrates the diversity of measures that can be used to investigate

transitions of consciousness, and emphasises that it is important to choose the right tools to answer the appropriate questions in every case.

1.8. CONCLUSIONS

This chapter started by surveying the problem of consciousness as it has existed for centuries, and framed it using perspectives from both neuroscience and philosophy. It then focused on transitions of consciousness and introduced the specific topics that will be addressed in this thesis – sleep, sedation and coma. Finally, several recent theories of consciousness were presented, which offer practical predictions regarding the relationships between neural activity and consciousness. This chapter provides a conceptual framework for the following studies presented in this thesis.

CHAPTER 2

COMPUTATIONAL METHODS

Given the theoretical framework discussed in Chapter 1, how can we measure consciousness in practice? This chapter critically surveys computational methods that can be applied to brain activity recordings to advance our understanding of neural processes related to consciousness. The measures described here are further used in the following chapters of this thesis. The focus is in particular on measures of spontaneous electroencephalographic (EEG) activity, but relevant findings are also drawn from other imaging techniques, such as fMRI. In line with the topic of this thesis, the main focus is on transitions between states, as opposed to contents, of consciousness. Three complementary approaches to quantifying brain activity are described: spectral measures, information-theoretical measures, and electric microstates of the brain.

2.1. THE ELECTROENCEPHALOGRAM

What can give us a clue about the fleeting events that occur in the minuscule cells tightly locked beneath our protective skulls? We currently have no tools that can measure non-invasively, in real time, the activity of individual neurons. Intracranial activity is occasionally studied in conjunction with brain surgery performed for reasons such as epilepsy (Téllez-Zenteno et al., 2005), but for research on large subject samples, an alternative window into brain functioning is needed.

A critical discovery in neuroscience has been that the electrical activity generated as neurons communicate produces a measurable electric field above the skull (Buzsáki et al., 2012). Taking advantage of this, the electroencephalogram (EEG) records the fluctuations of the electric field of the brain using electrodes placed on the scalp. In humans, this technique was pioneered and named by Hans Berger in 1924 (Haas, 2003). Models suggest that the EEG records postsynaptic potentials generated synchronously in cell assemblies (Niedermeyer and Lopes Da Silva, 2005). Source reconstruction algorithms can be used to estimate the location of brain sources underlying the EEG signal. In practice, however, source reconstruction is not always reliable (Hassan et al., 2014; Schoffelen and Gross, 2009). Markers computed directly on the signals recoded by individual scalp electrodes, termed the sensor space, can reliably provide signatures of cognitive states, including transitions between states of consciousness.

The EEG is not the only tool that can be used to measure brain activity. Alternative methods include magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and positron emission tomography (PET). Although these techniques have their own advantages, including potentially better spatial brain source resolution, they involve greater practical difficulties, such as lying still for a long time in the case of fMRI and MEG, and are considerably costlier. The EEG provides an easy and efficient solution that can be used at the bedside in clinical settings even with patients with impaired consciousness (Bagnato et al., 2010; Chennu et al., 2017; Chennu and Bekinschtein, 2012; Cruse et al., 2011; Fellinger et al., 2011; Harrison and Connolly, 2013). Moreover, along with MEG, the EEG provides a temporal resolution superior to other brain imaging methods, allowing the exploration of millisecond-level neural activity.

This thesis is focused on states of impaired consciousness. The EEG provides the best currently-available solution for investigating brain function during the process of falling asleep, entering sedation and, in particular, emerging from coma (Harrison and Connolly, 2013). These states will be investigated in the following chapters, with the aim to discover EEG signatures that can reliably distinguish, in real-time, between states of consciousness. This chapter continues by presenting the types of measurements performed on the EEG signal to reveal properties of neural activity, with a focus on changes already established to occur in sleep, sedation and coma, as relevant for the next chapters. The aim is to emphasise both the advantages and disadvantages of each individual tool in the greater picture of methodologies applicable to EEG data in the quest to understand the relationship between brain activity and consciousness.

2.2. SPECTRAL MEASURES

One of the most prominent characteristics of the human scalp EEG is the presence of oscillations with peaks at specific frequencies, which vary with the state of consciousness. The canonical frequency bands historically used to describe the EEG are denoted by Greek letters: alpha (8-13 Hz), beta (13-30 Hz), gamma (above 30 Hz), delta (below 4 Hz), theta (4-7 Hz). There are also other oscillations of interest in more specific contexts, such as sigma band for sleep spindles (12-15 Hz) or the mu rhythm (8-13 Hz) in the context of motor action. These intervals are only guidelines and can vary, particularly in experiments where altered or pathological states of consciousness are involved, such as sleep or disorders of consciousness. Each of these rhythms seems to bear a different cognitive meaning, although the exact source of each rhythm is not always fully clear and, sometimes, multiple unrelated sources produce rhythms of similar frequency which are difficult to untangle. Under several

consciousness frameworks, integration of information across the neural system plays an important role in the emergence of conscious states.

2.2.1. SPECTRAL POWER

During wakefulness, the alpha rhythm is a marker of relaxed wakefulness observed most prominently over the posterior, especially occipital, areas of the scalp. In healthy adults, eye closing results in bursts of alpha waves easily visible with the naked eye in the EEG (Barry et al., 2007). There is some debate on the physiology of the alpha rhythm, but its origin is often thought to be cortical (Niedermeyer, 2005a). As the subject becomes drowsy and approaches sleep, the alpha rhythm fades, while lower oscillations in theta and delta band appear in the EEG (Niedermeyer, 2005b; Ogilvie, 2001). Light sleep is characterized by spindles in the sigma frequency (12-15 Hz), along with other markers such as K-complexes. In deep sleep, slower frequency oscillations are present, including large-amplitude slow waves below 1 Hz, which reflect a cyclical hyperpolarisation and depolarisation of the membrane potential in cortical neurons (Steriade et al., 1993a, 1993c, 1993b), also referred to as up and down states (Wilson, 2008). Similarly to sleep, anaesthesia is characterised by the loss of alpha oscillations (Purdon et al., 2013) and the emergence of slower waves with similar cortical origins (Murphy et al., 2011). However, unlike sleep, anaesthesia is also often accompanied by the emergence of higher-frequency beta band oscillations (Purdon et al., 2013). By contrast, in pathological coma, spectral activity is disrupted depending on the aetiology of the damage of neural tissue (Kane et al., 1998; Lechinger et al., 2013). Therefore, different patterns of oscillatory activity in the EEG can provide useful indications for diagnostics and prognostication in a clinical setting, in conjunction with other biological and behavioural tests (Kaplan, 2004).

Some of the well-known oscillations are visible in the EEG with the naked eye. However, algorithms that produce a frequency decomposition of a time series can be used to obtain a comprehensive picture of the spectral architecture of the EEG. In this thesis, the Fourier transform and the Hilbert transform are employed to compute the spectral content of the data.

The discrete Fourier transform produces the representation of a signal from time domain to frequency domain by expressing it as a sum of sinusoids, thereby providing the power at *N* frequencies of interest:

$$S(f_k) = \sum_{i=0}^{N-1} s(t_i) e^{-j2\pi f_k t_i} (t_{i+1} - t_i), \qquad k \in \{0, 1, \dots, N-1\},$$

where s(t) is the waveform to be decomposed into a sum of sinusoids and S(f) is the Fourier transform of s(t). The fast Fourier transform provides a computationally efficient algorithm for this purpose

(Brigham, 1988). A downside of the Fourier transform is that it requires a window of data, assumed to be generated by a stationary and linear system, to compute the power spectrum, thereby producing a temporal resolution inferior to that of the original signal.

Alternatively, it is possible to use the Hilbert transform for better temporal resolution (Bendat and Piersol, 1986; Huang et al., 1998). The Hilbert transform of a real-valued function $x(t)$ $(where - \infty < t < \infty)$ is defined as the real-valued function:

$$H(x)(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(u)}{t-u} du$$

The Hilbert transform is the imaginary component of the analytic signal:

$$x_a = x(t) + j\, H(x)(t)$$

The analytic signal thus defines the instantaneous amplitude and phase of the original signal. By filtering the signal in the frequency band of interest, the instantaneous envelope obtained using the Hilbert transform can provide a useful spectral decomposition of the signal.

Historically, spectral power is the most well-established method of analysing the EEG and extensive literature exists on its relationship with a wide range of cognitive processes and states. Although some aspects are still debated, such as the origin of each individual rhythm, analysing the power spectrum is one of the most widely and reliable methods in the investigation conscious states. Different spectral configurations undoubtedly affect conscious processing (Hanslmayr et al., 2011; Klimesch, 2012) or necessarily accompany different conscious states (Massimini et al., 2007). But does this have any explanatory power for the emergence of consciousness? Given that consciousness and its loss involve coordinated changes across the whole brain network, it seems that the power spectrum by itself is not satisfactory in the quest for understanding consciousness. To understand how conscious states are sustained, a more global approach is needed, where the relationships between neuronal populations are considered.

2.2.2. SPECTRAL CONNECTIVITY

How can we quantify the relationships that form in a network consisting of 10^{11} neurons and around 10 times more glial cells (Bear et al., 1996)? To understand the processes that govern consciousness, which seems to encompass numerous cognitive functions implemented in different cortical locations, we ultimately need to understand the brain at both local and global level and how information flows and is integrated between different brain areas (Sporns et al., 2002). While anatomical connectivity is informative with regards to the infrastructure supporting neurophysiological dynamics, of more interest is functional connectivity, defined as the correlation between events occurring in remote

areas of the brain, as can be observed in live brain recordings (Fingelkurts et al., 2005). In particular, synchronisation of phase between signals originating from different brain sources is widely regarded as an indication of functional connectivity (Chennu et al., 2014) under the hypothesis that the synchronised firing of cell assemblies is fundamental in large-scale integration across neural systems in order to sustain coherent behaviour (Fell and Axmacher, 2011; Sauseng and Klimesch, 2008; Varela et al., 2001).

Brain connectivity can be investigated using the EEG by revealing mathematical relationships that suggest non-randomly synchronous activity between signals at different pairs of channels. In this investigation, we can use the same frequency bands as those described above for the power spectrum. At each point in time, an oscillation at a particular frequency can be described by its amplitude and its phase. A simple method of assessing linear correlations between signals is to use the coherence, computed as the cross-spectral density of two signals, which uses the Fourier transform of their cross-correlation (Sakkalis, 2011). However, coherence measures simultaneous changes in both amplitude and phase of two signals, whereas functional connectivity is manifested as a synchronisation of phase, but not necessarily of amplitude. To correct for this, indices based on signal phase only have been developed. For example, the phase locking value (PLV) (Lachaux et al., 1999) produces a value between 0 and 1 indicating the phase synchronisation between two signals. However, this measure has several disadvantages, one of which is volume conduction: the same source activity picked up by different electrodes is interpreted as true connectivity between different brain areas. The phase lag index (PLI) (Stam et al., 2007) was introduced to address the limitation of existing measures, such as the PLV, and to prevent the interpretation of volume conduction as true connectivity. Under the assumption that volume conduction results in either identical (0°) or opposite (180°) phases and a steady phase difference indicates connectivity, the PLI eliminates the former combination by averaging the signs of phase differences of the two signals. The PLI is computed using the imaginary part of the cross-spectrum X: $\Psi = |\langle sign(I(X)\rangle|$, where $\langle y\rangle$ denotes the expected value of y. Finally, the weighted phase lag index (WPLI) (Vinck et al., 2011) improves the PLI by weighting the signs of the phases by their absolute magnitudes, which corrects for the discontinuity of the measure and improves its sensitivity to noise:

$$\Phi = \frac{|\langle I(X)\rangle|}{\langle |I(X)|\rangle}$$

Considering the overall advantages of the WPLI over other methods, including the avoidance of volume conduction and the robustness to small sample sizes, the WPLI is employed in this thesis as a measure of connectivity. It should, however, be mentioned that the WPLI also has certain limitations. By down-weighting identical and opposite phases, it may discard true connectivity patterns. One study

has reported that the WPLI was successful in preventing the classification of volume conduction as true connectivity, but it underestimated connectivity in simulated data (Cohen, 2014). The WPLI can therefore be said to represent a conservative connectivity measure. Furthermore, the WPLI is not able to inform on the directionality of neural information flow; for this purpose, measures of directed functional connectivity should be used, such as transfer entropy (Schreiber, 2000) and Granger causality (Seth et al., 2015). Another family of methods that could be employed to measure effective connectivity is dynamic causal modelling (Friston et al., 2003).

The WPLI produces a list of numerical connectivity strengths between all pairs of nodes in the network. How can these connections be summarised to allow comparisons? A simple method is obtaining the mean of the median WPLI over regions of interest (Chennu et al., 2014). More elaborate metrics can be obtained using graph theory, which will be presented in section 2.4.

Particular resting-state connectivity patterns have been linked to different states of consciousness. A prominent marker consistently found in studies of impaired consciousness is long-range connectivity between frontal and parietal areas (Boly et al., 2013b). In healthy individuals, wakeful rest is characterised by frontoparietal connectivity at alpha frequencies (Chennu et al., 2016a, 2014), which is reduced in coma (Lehembre et al., 2012a), chronic disorders of consciousness (Chennu et al., 2014), sleep (De Gennaro et al., 2004) and anaesthesia (Ku et al., 2011). Other imaging methods, such as fMRI, have also confirmed that frontoparietal connectivity is altered across changing states of consciousness (Bor and Seth, 2012; Boveroux et al., 2010a; Heine et al., 2012). A recent study on patients with disorders of consciousness has confirmed that alpha connectivity was associated with metabolic activity in frontal and parietal brain areas in healthy adults (Chennu et al., 2017).

It is, however, still debated whether frontoparietal connectivity is a true signature of conscious processing (Bor and Seth, 2012; Naghavi and Nyberg, 2005) or merely a marker of a process which is difficult to disentangle from consciousness, such as goal-oriented tasks (Farooqui and Manly, 2017). In addition, there is currently lively debate regarding the content of consciousness and the location and nature of its neural correlates. Recent discussions have focused on the role of the posterior high-level sensory brain areas, including a hot zone comprising occipital, parietal and temporal areas, as a key region in generating conscious content (Boly et al., 2017; Koch et al., 2016). A proposed alternative is that the frontal cortex is necessary to support higher-level cognitive functions and consciousness (Odegaard et al., 2017). While these questions pertain to the content rather than the state of consciousness, they can nevertheless help us define better frameworks for studying its transitions, which include the subjective experience of falling asleep and that of emerging from traumatic brain injury.

Overall, spectral measures have been long-established in literature and are unquestionably valuable in practice for investigating cognition and conscious states. However, one disadvantage is that there exists no widely-accepted theoretical framework for consciousness that explains in a causal manner the relationship between different spectral components and consciousness. Although it has been observed, mainly through correlations, that distinct spectral profiles characterise different states of consciousness, no accepted explanation beyond association exists to provide insight into why particular frequency components should be linked to different states of consciousness. The next section described a set of alternative measures that can be more easily interpreted in the framework of modern theories of consciousness.

2.3. INFORMATION-THEORETICAL MEASURES

The concept of information integration as posited by consciousness frameworks, such as IIT, do not generally refer to the spectral content of the EEG. Instead, they refer to Shannon's definition of information (Shannon, 2001): the reduction of uncertainty caused by a specific outcome occurring from a set of possible outcomes. However, the difficulty in estimating states and outcomes in the neural system prevents a direct computation of measures of consciousness, such as the value of ϕ in IIT, the amount of information that the system can integrate to generate consciousness (Tononi et al., 2016). Instead, measures based on information theory can estimate the number of outcomes available to the system by examining the diversity of the signal within a time interval. Information exchange between different brain areas can also be estimated by quantifying the common patterns found in their respective signals.

The information-theoretical framework provides a window into a dimension of the EEG distinct from spectral measurements. Although communication between neural modules at specific biologically predefined frequencies means that information exchange will depend on these frequencies, the two approaches are conceptually separate and potentially orthogonal. Compared to spectral measures, the information-theoretical approach has a more direct foundation in theories of consciousness in complex systems. Such theories make qualitative predictions regarding the organisation and information flow in a system, but not necessarily on the specific frequency spectrum that should be fundamentally important in the emergence of consciousness.

2.3.1. SIGNAL COMPLEXITY

Current theories propose that, as we lose consciousness, the number of possible states in the neural system decreases (Tononi et al., 2016). In a system with a smaller number of possible states, uncertainty reduction in instantiating a conscious experience also becomes smaller. This also

potentially reduces the amount of information available for exchange across the brain. According to theories of consciousness, sleep, anaesthesia and coma should all have in common a quantifiable reduction in possible states (Seth et al., 2008; Tononi and Edelman, 1998), despite occurring due to different biochemical or mechanical causes. How can we quantify this loss of neural diversity from an information-theoretical point of view?

One method of assessing signal diversity makes use of the algorithmic complexity of a signal, as developed in computer science for file compression. The Lempel-Ziv algorithm and its variations (Lempel and Ziv, 1976; Welch, 1984) work by extracting a dictionary of unique patterns that appear in the input sequence and replacing these patterns in the sequence by their dictionary indices, thereby reducing the original size of the sequence. If the same patterns often repeat in the original sequence, the dictionary will contain a small number of items. On the other hand, if the sequence is composed of unpredictable, novel patterns, the dictionary will contain proportionately more items. Hence, the size of the dictionary quantifies the diversity of information contained in the sequence. By applying the same algorithm to a sequence of concatenated EEG data, the size of the dictionary can be interpreted as an approximation of the repertoire of states that the neural system displays. It has been shown that, in a sufficiently long signal produced by an ergodic process, the Lempel-Ziv complexity reflects the entropy of the process generating the signal (Schartner et al., 2017a).

Following early studies showing that the Lempel-Ziv complexity of the EEG tracks the depth of anaesthesia (Ferenets et al., 2007, 2006; Zhang et al., 2001), an influential application of this idea was the introduction of the perturbational complexity index (PCI) (Casali et al., 2013). To compute the PCI, transcranial magnetic stimulation (TMS) is applied to the resting-state EEG and source modelling is performed to obtain a binary matrix of significant and non-significant cortical sources. The temporal span of this matrix is on the order of hundreds of milliseconds. The Lempel-Ziv algorithm is then applied on this matrix to estimate its complexity. The PCI is defined as the Lempel-Ziv complexity c_L normalised using the length L of the matrix and its source entropy $H(L)$: $PCI = c_L \frac{\log_2 L}{LH(L)}$. The PCI was able to discriminate, in single individuals, between wakefulness and unconsciousness due to sleep or anaesthesia, and between levels of consciousness impairment in disorders of consciousness. In line with theories like IIT, this indicates that the neural response is significantly richer when subjects are conscious.

More recently, Lempel-Ziv complexity has been also applied to spontaneous EEG data. Although it is still designed to quantify the complexity of neural activity, the LZ_N measure is computed somewhat differently from the PCI. First, it is applied on non-perturbed EEG data. This data can be the raw EEG signal itself or its analytic signal. Secondly, the LZ_N is typically applied on data spanning a few seconds.

Finally, the normalisation process is different and will be detailed below. The Lempel-Ziv complexity of spontaneous EEG has been found to decrease in sleep (Schartner et al., 2017b), sedation and anaesthesia (Schartner et al., 2015), and disorders of consciousness (Wu et al., 2011). Interestingly, psychedelic drugs seem to enhance the complexity of the EEG (Schartner et al., 2017a), as also demonstrated in fMRI (Tagliazucchi et al., 2014).

To compute the Lempel-Ziv complexity of EEG data, the concatenation of a signal consisting of channel values over time can be performed either channel-by-channel or observation-by-observation, where an observation consists of the values of all channels at a single point in time. The interpretation of the two complexity flavours is slightly different: the former case reflects the local, temporal signal diversity in individual channel values over time, whereas the latter captures the spatial diversity of the global landscape of neural activity. In some of the above studies, a different flavour appears to have worked best in different contexts: for example, the spatial variant in anaesthesia (Schartner et al., 2015), and the temporal variant in psychedelic states (Schartner et al., 2017a). These different interpretations have not been thoroughly explored so far and it is not clear which variant best fits with the original theoretical framework that indicates neural information diversity as a key element for the emergence of consciousness. Bringing this investigation a step further, Chapter 4 presents evidence gathered from a sedation study on healthy adults that the two types of complexity track different aspects of sedation: spatial complexity tracks drug level, whereas temporal complexity tracks responsiveness.

Although currently not as well-established as the power spectrum, Lempel-Ziv complexity is a valuable alternative tool that is easy to compute and to deploy in clinical settings. It has the advantage of being more meaningful as a measure for the level of consciousness within theoretical frameworks that consider information diversity a key aspect of consciousness. Similar measures have confirmed that the loss of consciousness is characterised by lower complexity of neural activity (Bai et al., 2015; Hudetz et al., 2016). However, there are still questions to solve regarding the Lempel-Ziv complexity.

One ongoing debate regarding the Lempel-Ziv complexity is its relationship with spectral power. It is important to know whether changes in complexity can be explained, fully or partially, by changes in spectral power during different levels of consciousness. One way to address this question is to normalise the original complexity value by the maximal complexity of a sequence with the same power spectrum. This surrogate data for normalisation can be obtained by randomising the phases of the Fourier spectrum (Theiler et al., 1992). This normalisation was used to compute the LZ_N measure in other works (Schartner et al., 2017a). Alternatively, one very conservative approach is to apply a notch filter in order to remove a particular frequency from the signal completely, thereby ensuring that

frequency band has no effect on the result. Chapter 4 of this thesis uses both phase normalisation and notch filters to confirm that Lempel-Ziv complexity is not explained away by changes in power.

2.3.2. INFORMATION SHARING

In addition to quantifying the repertoire of states available to the neural system, we can attempt, under the same consciousness frameworks, to measure how information is integrated, or shared, across the system. Similarly to spectral analysis, by observing patterns originating from different brain areas that occur synchronously, a pairwise connectivity map can be created. Just like in the case of phase relationships, identical patterns can be regarded as volume conduction instead of indicators of connectivity.

One measure that has been introduced for this purpose is the weighted symbolic mutual information (wSMI) index (King et al., 2013). The wSMI is computed by transforming the EEG signal into symbols and computes their joint probability of occurrence between each pair of channels. The transformation is performed as follows. First, a temporal separation parameter τ is chosen, that defines the distance between the selected EEG samples. Then, k values (e.g. $k = 3$) are selected given this distance. This sequence of length k is labelled with a unique symbol that depends only on the relative ordering of the values in the sequence. The joint probability for each pair of symbols is then computed. Importantly, symbols with identical or opposite shapes are considered to be generated by the same sources and ignored, hence correcting for volume conduction. The symbols are thus able to capture nonlinear coupling between signals. The sensitivity to particular frequency bands can be tuned by adjusting the temporal separation τ between the samples that contribute to a symbol. The wSMI has been shown to decrease in disorders of consciousness in comparison to healthy wakefulness (King et al., 2013; Sitt et al., 2014), particularly at theta frequencies, although the differences do not always reach significance (Claassen et al., 2016). A recently introduced alternative to the wSMI is the phase lag entropy (PLE) (Lee et al., 2017), which analyses the patterns present in phase relationships between two signals to provide a diversity of temporal patterns of functional connectivity. The PLE reliably distinguished between consciousness and its loss in anaesthesia (Lee et al., 2017).

The wSMI is currently less established in literature compared to many other methods of assessing connectivity. However, it has the advantage of being a measure directly inspired from theories of consciousness, as opposed to being constructed in a data-driven manner. Existing studies offer moderate evidence that the wSMI could potentially be useful in measuring integration in the neural system and hence this measure is used in Chapter 4 as a complement to signal complexity.

After obtaining a full matrix of connectivity values employing the wSMI, the same question applies as in the case of spectral connectivity: how to best make sense of the large number of connectivity values describing the network of information-sharing? In the next section, graph theory will be presented as one method of delving into the topographical details of connectivity networks.

2.4. GRAPH-THEORETICAL MEASURES

Modern theories of consciousness support the view that global patterns of information exchange, rather than activity in specific biological locations, account for the emergence of consciousness (Tononi et al., 2016). Integration across large scales in systems consisting of a large number of states is hypothesised to play an important role in this process. While exact connections between large numbers of cells in the neural system are currently impossible to compute in vivo, brain imaging and recording techniques such as the EEG can be used to investigate the activity of populations of neurons. As described in sections 2.2.2 and 2.3.2, connectivity methods applied to the EEG provide us with a manageable network up to the order of a hundred nodes, where emerging patterns can be easily discovered.

A mathematical tool that can be applied to such networks is graph theory (Bullmore and Sporns, 2009). Given a network consisting of nodes connected by edges of different values, which define a graph, its structure can be summarised using properties at local or global scales. In particular, it has been proposed that human brain connectivity has a small-world architecture: dense short-range and few long-range connections, which result in a very low path between any two nodes in the network, whether topologically close or far from each other (Bassett and Bullmore, 2006, 2016; Sporns, 2010). This architecture allows both segregated and integrated information processing in the brain to occur with high efficiency (Achard and Bullmore, 2007). In fact, small-world architectures are encountered in many other natural and artificial complex systems, from human social circles (Milgram, 1967) to protein structures (Taylor, 2013).

The two essential properties which define a small-world architecture – segregation and integration – are measured with different metrics (Watts and Strogatz, 1998), most of which can be applied on both directed and undirected graphs. To quantify the potential for segregated processing, it is useful to analyse local, direct connections between nodes, and assess the density of local connections using measures such as the node degree or the degree of clustering. On the other hand, the integration potential of the network is best given by global measures of distance, such as the average path length, between pairs of nodes, which includes node pairs that are topographically far and not necessarily linked by direct connections. Small-world-ness can be calculated as a ratio between segregation and

integration properties of the network (Humphries and Gurney, 2008). In addition, certain nodes, corresponding to highly connected biological areas, play a special role within the network and act as hubs: they form a higher number of long-range connections, like railway stations and airports that connect distant regions. Moreover, networks can be separated into modules: separate groups of densely-interconnected nodes which can be assumed work together to process information (Guimera et al., 2004; Newman and Girvan, 2004). The distance covered by such modules can inform on the long-distance processing occurring in the network, as measured, for example, by the modular span (Chennu et al., 2014).

In the study of conscious states, graph theory builds on connectivity methods to allow a more detailed survey of functional network architecture. Several studies have made use of this technique. In chronic disorders of consciousness, the EEG shows that alpha networks have a smaller number of hubs, less efficiency and cover smaller brain areas in chronic disorders of consciousness compared to healthy adults, whereas lower-frequency theta and delta networks show the reverse modifications (Chennu et al., 2014). In comatose patients, BOLD networks show a number of preserved properties, but hub structure is disrupted (Achard et al., 2012). In propofol anaesthesia, hubs are also reorganised (H. Lee et al., 2013) and the average path length is increased, suggesting that loss of global integration is associated with unconsciousness (Monti et al., 2013). Chapter 5 will show how graph-theoretical measures that quantify the early changes in connectivity network properties weeks after traumatic brain injury can assist in prognosticating the eventual outcome of acute comatose patients.

A downside of applying graph theory to brain networks is that multiple levels of abstraction and indirection are involved. First, EEG records activity over surfaces of the scalp, which approximates activity in a particular brain region. Secondly, connectivity is inferred using methods which are reliable, but have limitations, as emphasised in the previous sections. Finally, to apply graph theory, connectivity networks are usually further modified in order to allow well-defined computations. For example, where the graph is defined by edges of strengths between 0 and 1, a common option is applying a threshold to retain only the strongest connections (Reijneveld et al., 2007). The remaining connections can then be set to 1 to obtain an unweighted graph, which simplifies the computation of graph-theoretical properties that operate with binary edges, such as the clustering coefficient or modularity measures. However, this discards information present in the weights, which other measures can take advantage of. But what is an optimal threshold that best delimits strong from weak connections? There is no current consensus on this: one option is to perform the same computation on a range of thresholds and average the results (Achard et al., 2012; Chennu et al., 2014; Lynall et al., 2010). However, this might have a smoothing effect on the results. Moreover, the values of the selected thresholds are not consistent across literature.

To sum up, graph theory supplements connectivity measures by providing tools for more detailed analyses of network architecture, thereby revealing global and local properties which characterise the integration and processing of information in the brain. These patterns are both informative as signatures of clinical alterations of states of consciousness and as pointers in current theories of consciousness. However, a high degree of abstraction is involved in this process, so any findings should be taken with caution.

2.5. ELECTRIC MICROSTATES

The above measures seek to quantify local neural activity, as well as the relationships between activity in different brain regions. In contrast, the concept of electric microstates of the brain starts from a different perspective. It refers to the momentary global state of the brain at a specific point in time and the parameters of the sequence produced by the rapid succession of these global states. One commonly-used method to establish a global state is the topography of the electric field of the brain (Michel et al., 2009). A finite number of such quasi-stable global states with a duration in the range of tens of milliseconds – hence termed *microstates* – have been consistently found to occur in healthy humans (Koenig et al., 2002) and psychiatric disorders (Strelets et al., 2003; Tomescu et al., 2014), across different cognitive states (Milz et al., 2015). Crucially, four canonical states denoted by letters from A to D typically occur in the classic EEG microstates paradigm (Figure 2.1). It is not the states that change across conditions, but their dynamics, such as the duration or the relative frequency of particular microstates.

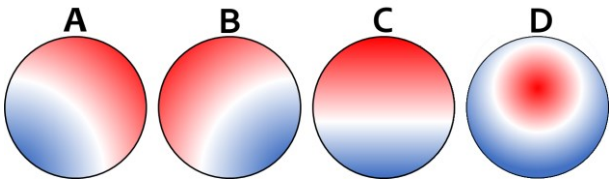

Figure 2.1 ***Illustration of the four typical microstate topographies consistently found in literature.***

The classic algorithm for finding microstates in EEG data involves an unsupervised clustering process (Michel et al., 2009; Murray et al., 2008; Pasqual-Marqui et al., 1995). A sample consists of the map of electric voltage values from all selected EEG channels at a single point in time, which define a static topography on the scalp. The algorithm clusters the given EEG samples heuristically and produces a fixed number of topographies (maps) that best approximate the most frequently occurring microstates in the given set of samples. Usually, the input set consists of the samples with local maximum variance, which provide the best-defined topographies across the data. Then, the resulting topographies are back-fitted to the original data by assigning to each sample the microstate map with

the highest spatial correlation. Hence, the original sequence of EEG samples will become a sequence of microstate topography labels. Contiguous periods with a common label represent a quasi-stable microstate. The average duration (number of contiguous samples) of each microstate can be used as a parameter that characterises the sequence of global brain changes, along with other parameters such as the frequency of each microstate occurring in a period of time, the goodness-of-fit of the samples to the fixed microstate maps, or the transition matrix between microstates.

Although by now well-established and widely used in literature, the methodology has several implementation details which vary across studies. First, the clustering algorithm produces a predefined number of microstate maps. How to decide on the number of microstates that best explain the data? More microstate maps will always explain more, but a balance between their number and the variance explained needs to be found. A cross-validation criterion has been proposed for this purpose (Pasqual-Marqui et al., 1995). However, this criterion has been reported to be too easily influenced by the number of electrodes used in the algorithm(Murray et al., 2008), a finding confirmed in Chapter 3. Using this criterion (Brodbeck et al., 2012; Kuhn et al., 2015; Van de Ville et al., 2010) or other variance-related criteria (Koenig et al., 1999) some studies have found four microstate maps to be optimal in most (although not all) cases. However, many more other studies use an a priori number of four microstates based on previous studies, e.g. (Khanna et al., 2014; Kikuchi et al., 2011; Koenig et al., 2002; Milz et al., 2015; Schlegel et al., 2012; Tomescu et al., 2014). While this allows a comparison between the functional roles of individual microstates, it also restricts the study to a predefined set of microstates which still do not have a clear meaning. Furthermore, different methods of segmenting the EEG into microstates may result in a different optimal number of maps (Yuan et al., 2012).

Secondly, there are two ways of performing the back-fitting procedure. The first method (Milz et al., 2015) makes the assumption that topographies change in between local peaks of variance (global field power local maxima) and hence labels only these peaks based on their spatial correlation with the maps obtained through clustering. The labels of the EEG samples in between peaks are interpolated. While this avoids noisy assignments, it also potentially discards shorter microstates occurring between peaks and causes a longer erroneous assignment if a single peak is assigned to the wrong microstate. In contrast, the second method (Tomescu et al., 2014) uses more computational power to label every single EEG sample using the microstate maps obtained by clustering. This generates a more detailed impression of the sequence of microstates occurring throughout the recording, but it is more prone to noise. For this reason, a smoothing algorithm can be applied to correct isolated labels different from their neighbours (Pasqual-Marqui et al., 1995). However, even after smoothing, noise can still be present in the final result, especially given that samples situated between peaks of global field power may not have well-defined topographies.

Finally, the neuroanatomical and functional interpretation of microstates is not straightforward, and the relationship between microstates and other brain measures, such as spectral power and connectivity, is currently unclear. A combined fMRI-EEG study suggested that each microstate corresponds to four resting-state networks linked to auditory, visual, attention and interoceptive processing, but none to the default mode network (Britz et al., 2014). It also found no correlation between microstates and the power spectrum. On the other hand, a source localisation study identified individual microstates as parts of the default mode network, proposing that the default mode network is in fact the sum of fast-paced individual components whose separation is smoothed by the low temporal resolution of fMRI (Pascual-Marqui et al., 2014). A further study reported that different alpha oscillation sources are correlated with each microstate, indicating an inhibitory sequence in the default mode network (Milz et al., 2017). More work is needed to clarify the significance of these results. Moreover, the relationship between microstates and brain connectivity or complexity is also an open question. Current evidence suggests that neural activity can indeed adapt at sub-second scales to produce distinct stable spectral power and connectivity patterns (Vidaurre et al., 2016), so the relationship between these and EEG microstates is an interesting exploration avenue. Chapter 3 investigates this question to reveal a previously unknown relationship between a particular microstate and fast connectivity patterns at the onset of sleep.

Overall, EEG microstates have revealed interesting properties of brain activity. When the microstates were first described, they were proposed to be the building blocks of cognitive processes (Lehmann, 1971). A range of healthy and pathologically altered cognitive states have been described using EEG microstates. An example is mental disorders: in schizophrenia, a shortening of microstate D has consistently been found (Kikuchi et al., 2007; Lehmann et al., 2005; Nishida et al., 2013), while microstate C has been found to be altered in dementia (Grieder et al., 2016; Nishida et al., 2013). Other microstate alterations have been linked to different cognitive modalities and processes (Milz et al., 2015; Seitzman et al., 2016). During transitions of consciousness such as sleep (Brodbeck et al., 2012) and hypnosis (Katayama et al., 2007), the parameters of the EEG microstates appear to change, but the microstate topographies do not.

How can microstates improve our current understanding of consciousness? While most theories of consciousness address the spatial aspect of neural activity strength, connectivity and diversity, microstates provide a fine-grained lens into the temporal dynamics of conscious states. One transition of consciousness where EEG microstates can provide insight is the process of falling asleep. As we become drowsy, how do the rapid microstate dynamics change as a function of being conscious or unconscious? In Chapter 3 of this thesis, this question is addressed by investigating the microstate dynamics that accompany the loss of responsiveness during the onset of sleep. In answering this

question, a previously unknown link will be presented between EEG microstates and the changes already established to occur in sleep using the complementary methods of spectral power and connectivity.

2.6. CONCLUSIONS

This chapter introduced three complementary methods of analysing brain activity using EEG data: spectral analysis, information-theoretical analysis, and electric microstates. Spectral and information-theoretical analyses can be useful for investigating two complementary aspects of brain function: activity recorded from individual neuronal populations, and the pairwise connectivity between these populations. While spectral analysis investigates the distribution of power into the oscillatory components of the signal and their phase relationships, information-theoretical approaches transform the signal into symbols and quantify their diversity and co-occurrence. These approaches are suitable for testing predictions of current theories of consciousness regarding a reduced amount of diversity and integration of information in the neural system during natural, pharmacological and pathological unconsciousness, as opposed to healthy wakefulness. On the other hand, electric microstates describe a sequence of momentary, global states of the brain. This sequence has been shown to slow down as people fall asleep and particular topographies have been associated with altered cognitive states. The microstate sequence can offer insight into the temporal diversity of the repertoire of brain states. Overall, these methods all provide promising avenues for ascertaining the neural correlates of consciousness and its loss. Importantly, no single method can be identified as superior to the others. They can all be used in a complementary manner, taking into consideration their respective advantages, weaknesses and meaningfulness in the context provided by relevant theories of consciousness.

CHAPTER 3

SLEEP

How do we lose consciousness as we fall asleep? This chapter investigates this question under the theoretical and methodological frameworks developed in the previous chapters. For this study, responsiveness to a simple auditory categorisation task is used as a proxy to zoom into the gradual loss of consciousness between wakefulness and sleep, keeping in mind the limitations of this method as discussed in section 1.5. This chapter analyses the neural information integration supported by long-range frontoparietal connectivity observed during wakefulness in alpha band and shows that it breaks down during unresponsiveness, while connectivity at theta frequencies emerges between the same regions. Further, it is shown that the temporal dynamics of rapidly-changing EEG microstates slow down as participants stop responding. A specific microstate (D) is identified whose increased duration predicts unresponsiveness at single trial level. Finally, combining for the first time these two methods, a novel relationship between microstates and brain networks is exposed, as it is found that microstate D uniquely indexes significantly stronger theta connectivity during unresponsiveness. These findings suggest that the transition to unconsciousness is not linear, but rather consists of an interplay between transient brain networks reflecting different degrees of sleep depth.

Parts of this chapter have been presented as a poster at the International Conference for Cognitive Neuroscience (ICON) 2017 in Amsterdam. This chapter forms the basis for the article published in the journal Brain Topography (Comsa et al., 2018), available online at http://dx.doi.org/10.1007/s10548-018-0689-9. The code used in the analyses described in this chapter can be found at https://github.com/iulia-m-comsa/EEG/tree/master/Microstates. The data is available at https://doi.org/10.17863/CAM.33597.

3.1. INTRODUCTION

As we fall asleep, our brain traverses a series of changes which accompany the loss of sensory awareness and responsiveness to the external world. Despite the subjective ability to classify retrospectively one's own state as ‘awake’ or ‘asleep’ (Hori et al., 1994), research continues to unravel the gradual transitions happening at behavioural (Ogilvie and Wilkinson, 1984), cellular (Steriade et al., 1993a), physiological (Prerau et al., 2014) and cognitive (Goupil and Bekinschtein, 2012) level,

starting with early drowsiness and continuing into the deep stages of sleep (Ogilvie, 2001). Characterising these transitions and linking across physiological levels is an important step in the modern attempt to understand access-consciousness (Block, 1996; Koch et al., 2016) and its fluctuations in natural, pathological and pharmacological alterations: sleep (Hobson and Pace-Schott, 2002), disorders of consciousness (Giacino et al., 2014), sedation and anaesthesia (Alkire et al., 2008).

The transition from wakefulness to sleep involves a progressive and sometimes nonlinear loss of responsiveness to external stimuli (Ogilvie and Wilkinson, 1984). Behavioural unresponsiveness does not immediately imply unconsciousness (Overgaard and Overgaard, 2011; Sanders et al., 2012). However, from the perspective of levels of consciousness (Laureys, 2005), the capacity to respond to external stimuli offers an objective measurement in the process of transition between full wakefulness and sleep-induced unconsciousness. The question of how we stop responding to stimuli during drowsiness is related to, but distinct from an investigation of the stages of sleep conventionally defined by specific electrophysiological grapho-elements such as K-complexes and sharp waves, or the variation in spectral power at slow frequencies (Iber et al., 2007; Ogilvie, 2001). Indeed, the loss of responsiveness is distributed across sleep stages: one study found a rate of unresponsiveness of 28% in stage 1, 76% in stage 2, and 95% in stage 3 of sleep (Ogilvie and Wilkinson, 1984). Here, we are specifically interested in the neural markers that predict our inability to respond as we drift to sleep.

A traditional approach for investigating this question is to look at the changes in EEG spectral power and connectivity, which have been shown to vary across levels of consciousness. During relaxed wakefulness, the EEG of most human subjects is characterised by trains of alpha waves, at around 10 Hz, originating from central-posterior cortical areas (Barry et al., 2007; De Gennaro et al., 2016; Niedermeyer, 2005a). During the early onset of sleep, these alpha oscillations disappear and an alpha rhythm with a different cortical origin (Broughton and Hasan, 1995) emerges in anterior regions (Tanaka et al., 1997), while theta power increases, particularly in central regions (Badia et al., 1994; Niedermeyer, 2005b; Ogilvie, 2001; Wright et al., 1995). Similarly, long-range alpha connectivity disintegrates at the onset of sleep, while lower-frequency theta and delta connectivity increases (Tanaka et al., 2000, 1998; Wright et al., 1995). Several power and connectivity patterns have been associated with the loss of consciousness, sometimes specifically with the loss of responsiveness, such as the anteriorisation of alpha power and connectivity in EEG, which has been described in drug-induced loss of responsiveness (Chennu et al., 2016a), and frontoparietal connectivity, which has been proposed as a key signature of consciousness (Bor and Seth, 2012; Laureys and Schiff, 2012) and linked to external awareness (Vanhaudenhuyse et al., 2011). The disruption of frontoparietal connectivity at alpha (8-12 Hz) frequencies has been shown to occur in disorders of consciousness (Chennu et al., 2014) and sedation (Chennu et al., 2016a). Although it is still debated whether these are signatures of

conscious processing or of processes that almost invariably accompany it (Farooqui and Manly, 2017), brain connectivity patterns currently provide, in practice, useful measures in the context of transitions between levels of consciousness.

Another method that can be employed to investigate the rapidly changing global state of the brain is that of EEG microstates. A microstate represents a quasi-stable spatial topography of electric field on the scalp (Lehmann, 1990, 1971; Lehmann et al., 1987). The conventional method of analysing microstates in a dataset involves running an unsupervised clustering algorithm on a set of EEG topographies of highest variance, followed by labelling of all EEG samples based on the similarity with the four obtained topographies (Murray et al., 2008; Pasqual-Marqui et al., 1995). Four consistent (Khanna et al., 2014) EEG microstate topographies have been identified in a large population of healthy subjects of all ages during resting-state wakefulness (Koenig et al., 2002) and different microstates have been correlated with different cognitive modalities (Lehmann et al., 2010; Milz et al., 2015; Seitzman et al., 2016), but also with mental disorders, such as narcolepsy (Kuhn et al., 2015). A resting-state study of sleep (Brodbeck et al., 2012) identified four EEG microstate topographies in all stages of sleep nearly identical to those of wakefulness, but occurring with altered temporal parameters. Notably, increased microstate duration was associated with deeper sleep. On the contrary, a different study (Cantero et al., 1999) reported a shorter duration of microstates and suggested a larger repertoire of brain states during the hypnagogic period. Microstates are thought to reflect momentary, global, synchronised (Koenig et al., 2005) networks of the brain, reflecting building blocks of large-scale cognitive processing required for the continuous stream of consciousness (Lehmann, 1990). The neural sources underlying microstates are still being explored (Britz et al., 2010; Milz et al., 2017; Pascual-Marqui et al., 2014). Still, the dynamics of the sequence of microstates itself can be seen as a 'syntax' of neural activity that is in and of itself an informative tool for modelling and understanding the rapidly-fluctuating global dynamics of the brain.

Brain connectivity and microstates hence provide complementary perspectives on the neurodynamics underlying the loss of responsiveness as we fall asleep. But what is the relationship between brain networks and microstates? There is evidence that transient brain networks can be resolved in electrophysiological data (Baker et al., 2014; Pascual-Marqui et al., 2014; Vidaurre et al., 2016), but it is an open question whether these networks co-occur with the lifetime of individual microstates. We investigate for the first time how spectral connectivity and EEG microstate dynamics interact as we lose responsiveness during drowsiness. We hypothesise that the spectral changes occurring with the loss of responsiveness mirror those observed in the transition to sleep (Ogilvie, 2001), anaesthesia (Chennu et al., 2016a; Purdon et al., 2013) and in disorders of consciousness (Chennu et al., 2014): namely, the disintegration of alpha networks, the loss of posterior alpha power, and the emergence

of lower-frequency connectivity and power. Alongside, building on previous research on EEG microstate dynamics during sleep (Brodbeck et al., 2012), we hypothesise similar changes in microstate dynamics accompanying the loss of responsiveness during drowsiness. Finally, given that resting-state network activity is known to fluctuate at millisecond level, we hypothesise that the neural changes that occur during drowsiness underlie the dynamics of both brain networks and the microstates sequence. Specifically, we investigate the possibility that individual microstates co-occur with distinct transient brain networks, reflecting fleeting changes in the global state of the brain during drowsiness.

To address these questions, we use a subset of data from a previously reported auditory discrimination task where subjects became drowsy and unresponsive (Kouider et al., 2014). The task involved pressing a button corresponding to the classification of the auditory stimulus into one of two categories (object or animal). We obtain five minutes of data as subjects performed this task, before and after the loss of responsiveness due to drowsiness. We first characterise the responsive and unresponsive periods by analysing microstate-blind spectral power and connectivity changes in our dataset. Next, we describe the temporal parameters of EEG microstates during responsiveness and unresponsiveness. To test whether these parameters can reliably predict responsiveness to individual stimuli, we apply machine learning to predict responses and misses to stimuli in our task, based only on pre-stimulus microstate parameters. Finally, we investigate the brain connectivity underlying each of the four canonical microstates after the loss of responsiveness and highlight a previously unknown relationship between spectral connectivity and EEG microstates.

3.2. METHODS

3.2.1. SUBJECTS

Sixteen healthy, native English-speaking, right-handed young adults (mean age = 24, S.D. = 2.75; 6 females) were selected for this experiment out of the eighteen subjects from Experiment 1 in a previous study (Kouider et al., 2014). Two subjects from this dataset were excluded by visual inspection due to a failure to remain asleep for a period longer than five minutes, as assessed using responsiveness to stimuli. The participants were directed to not consume stimulants like coffee and to sleep 1-2 hours less than normally before the experiment. All of the subjects were assessed as easy sleepers on the Epworth Sleepiness Scale (scores 7-14). The participants signed a consent form and were reimbursed for their participation. The experiment was approved by the Cambridge Psychology Research Ethics Committee.

3.2.2. EXPERIMENTAL PROCEDURE

The stimuli consisted of 96 spoken English words chosen from the CELEX lexical database (Linguistic Data Consortium, University of Pennsylvania). Half of the words denoted animals and the other half denoted objects. The subjects were asked to classify each stimulus in its respective category (animal or object) by pressing a button. The stimuli were presented through headphones, with an average distance of 8.4 seconds (minimum 6.2 seconds) between consecutive stimuli, as the subjects were lying with their eyes closed in a reclining chair. To facilitate drowsiness, the task was performed in a dark, acoustically and electrically shielded EEG room, and the participants were told that they could fall asleep at any point during the experiment, although they were asked not to stop responding deliberately while still awake.

3.2.3. EEG DATA ACQUISITION

The electroencephalogram was continuously recorded at 500 samples per second from 64 Ag/AgCl electrodes (NeuroScan Labs system) positioned and labelled according to the extended 10/20 system, with Cz as a reference and including vertical and horizontal electrooculography channels.

3.2.4. EEG PRE-PROCESSING

All analyses that follow were performed using custom MATLAB scripts (The MathWorks, Inc., Natick, Massachusetts, US). The EEGLAB toolbox (Delorme and Makeig, 2004) was used to facilitate data pre-processing.

The data was filtered between 1 and 40 Hz and the full channel mean was subtracted from each channel for baseline correction. The HEOG and VEOG channels were removed. An Independent Component Analysis (ICA) decomposition was performed using the infomax ICA algorithm (Bell and Sejnowski, 1995). Components capturing ocular or single-channel artefacts were removed from the data by visual inspection and considering the correlation with the HEOG and VEOG channels. An average of 11.6 (S.D. = 8.6) out of 63 components were removed per subject. Channel FT8 was interpolated using spherical interpolation in all subjects due to being noisy in most recordings. Finally, channels were re-referenced offline to the common average.

3.2.5. DATA SEGMENTATION

We classified responsive and unresponsive periods by inspecting the sequence of hits and misses to individual stimuli. We used a liberal window of 6 seconds to allow for a response to a stimulus, regardless of its correctness. A lack of response within 6 seconds was marked as a miss. The choice of a 6-second window for responsiveness was based on our own pilot studies, where we investigated the

longest interval that subjects would make a response during drowsiness in a go task. However, note that most reaction times were below 3 seconds (Figure 3.1) and the reaction times increased gradually and later in the task, indicating an increase in drowsiness. This was also established in a previous study on the same data (Kouider et al., 2014).

For balance across participants and the two behavioural states, a total of five minutes of responsiveness and five minutes of unresponsiveness were extracted from each recording (150000 samples per state, per recording), as shown in Figure 3.1. The responsiveness period was taken as the first 0.5 to 5.5 minutes of data in each recording, acquired immediately after the experiment began and the participants were still alert and wakeful. This was confirmed by checking that the large majority of the stimuli were followed by responses during this period; a very small number of occasional misses occurred in more than half of the participants during this period (e.g. to the unfamiliarity with the task), but they were not contiguous. Then, a period of unresponsiveness was selected by visual inspection of the hits and misses after the end of the responsiveness period, with the aim to find a five-minute interval consisting of as many misses as possible. If a response was present during the period labelled as unresponsiveness, the 10 seconds preceding and following the corresponding stimulus were excluded.

3.2.6. MICROSTATE TOPOGRAPHIES

The computation of the sequence of EEG microstates is based on the observation that the topography of the electric field recorded by EEG over the scalp does not fluctuate randomly, but rather comprises short periods of stability (Lehmann, 1971). To compute the microstate topographies, the global field power (GFP), representing the standard deviation of the electrode values (Lehmann and Skrandies, 1980), was first computed at each time point. As the number of GFP peaks varied across subjects and condition, we rounded down the minimum number of peaks available and retained the first 5000 peaks in each condition (responsiveness and unresponsiveness) from each recording.

The clustering algorithm was implemented in MATLAB and is presented in Box 3.1. The algorithm is based on a variant of the method first introduced by (Lehmann et al., 1987), as described in (Murray et al., 2008), and involves an unsupervised clustering of EEG samples into the specified number of classes that best explain the input samples. Note that topographical similarity is computed using the absolute value of the spatial correlation and the polarity of the map is ignored, as topographies with inverted polarities are considered to be produced by the same neural generators (Michel et al., 2009). The maximum number of iterations was set to 1000 and the GEV delta was set to 1e-9.

We initially employed a cross-validation criterion (Pasqual-Marqui et al., 1995) to determine the optimal number of microstates fitting the data, as performed in several previous studies (Brodbeck et al., 2012; Koenig et al., 1999). However, we found that the cross-validation criterion produced different results for when the number of electrodes was down-sampled from 63 to 30 (7 and 4 maps, respectively). This sensitivity of the cross-validation criterion to the number of electrodes has been documented in previous literature (Murray et al., 2008). Hence, we decided to fix the number of microstates to four, similarly to previous studies that also fix this number a priori (Khanna et al., 2014; Kikuchi et al., 2007; Koenig et al., 2002; Milz et al., 2015; Schlegel et al., 2012; Strelets et al., 2003; Tomescu et al., 2014).

The scripts used for generating the microstate maps and computing the parameters of their sequence, as described in the following sections, are available online at https://github.com/iulia-m-comsa/EEG.

Microstate clustering algorithm

Input: *n* average-referenced EEG samples (*n* x *number_of_channels*) from GFP peaks.

Output: *k* maps that best characterise the data.

1. Normalise each input sample to a vector of length 1.
2. Pick k random samples as the initial maps.
3. Label each sample as $i \in \{1, \ldots k\}$, where *i* is the index of the map with the highest absolute spatial correlation.
4. Re-compute each map *i* as the first principal component of each cluster of samples labelled *i*.
5. Compute the Global Explained Variance (GEV).
6. If GEV delta is small enough or maximum number of iterations has been reached, end. Else, go to 3.

Box 3.1 ***Microstate clustering algorithm.***

3.2.7. MICROSTATE LABELLING

To obtain the sequence of EEG microstates present during a recording, each EEG sample was individually assigned to the microstate with the highest corresponding spatial correlation. To correct for noisy assignments during polarity reversals (Koenig and Brandeis, 2016), we applied a previously-described temporal smoothing algorithm for the microstate sequence (Pasqual-Marqui et al., 1995) with the half-width of the smoothing window (parameter b in Pasqual-Marqui's algorithm) set to 5

samples, corresponding to a smoothing neighbourhood of 20ms. This parameter was chosen to be in the range of mean microstate durations found by (Gärtner et al., 2015) using a model of microstate transition processes based on Markov chains (10 ms during wake, 34 ms during deep sleep).

3.2.8. MICROSTATE PARAMETERS

Following the full labelling of each recording, three parameters were computed for each microstate per state (responsiveness and unresponsiveness) and per recording:

- The *microstate temporal coverage*, also called the *fractional occupancy*, indicating the percentage of time spent in one microstate;
- The *microstate duration*, indicating the average length of continuous sequences labelled as one microstate;
- The *global explained variance (GEV)*, representing the amount of spatial correlation of the samples with their corresponding microstate topography, normalised by the GFP.

3.2.9. STATISTICS

Interactions between microstate parameters and behavioural state (responsiveness and unresponsiveness) were performed using a two-way repeated measures ANOVA (Hogg and Ledolter, 1987) with the microstate label and the behavioural state as factors. Sphericity was tested using Mauchly's test (Mauchly, 1940) and, where violated, was corrected using the Greenhouse-Geisser procedure (Greenhouse and Geisser, 1959). The Tukey-Kramer method (Tukey, 1949) was used to correct for multiple comparisons. After correction, a conventional threshold of $p=0.05$ was used to assess significance. Unless otherwise specified, similar statistical tests were also performed for the measures that follow.

3.2.10. RESPONSIVENESS PREDICTION

We applied machine learning classification to explore whether microstate properties identified in the ongoing brain dynamics immediately preceding each auditory stimulus in the experimental trials could predict the presence or absence of a response to that stimulus. Importantly, all trials were considered for classification, both within and outside the periods labelled as responsive or unresponsive for the above microstate analysis.

Five seconds of EEG data immediately preceding a stimulus were used to generate the features for classification. We also investigated using shorter pre-stimulus time periods, down to 1 second of pre-stimulus data, but we found that classification accuracy increased with a larger amount of pre-stimulus data over which microstate dynamics could be more accurately estimated. At the same time, the

amount of pre-stimulus data was restricted by the overlap with the previous trial. Trials overlapping with a response corresponding to the previous stimulus were excluded. By setting the pre-stimulus window to five seconds, less than 10% of the trials were rejected due to overlap with the previous trial.

The input features generated for classification consisted of either individual microstate parameters computed during the five-second pre-stimulus period in each trial, or a combination of these parameters. The parameters were those we previously characterised at the group level: namely the mean duration, mean coverage, and mean GEV for each microstate separately. The classifier was trained separately with the above individual and combined features. As a baseline, the theta-alpha ratio was also computed for each trial as the ratio between the total power spectral density at 5-6 and 9.5-10.5 Hz respectively, and used as an input feature for the classifier. The classification label for each trial was generated by labelling it as either as a timely response (1) or a miss (0).

We employed leave-one-subject-out cross-validation to test for the generalisability of the classifier's performance. For this, the data was split into 16 folds, with one fold corresponding to a single participant's trials. A support vector machine (SVM) (Christianini and Shawe-Taylor, 2000) with a radial basis function kernel (Vert et al., 2004) was trained repeatedly by excluding one fold at the time from the training set and using it as a test set. The SVM was optimised by exhaustive search to use the optimal value for two parameters: the box constraint, which restricts the number of support vectors, and the kernel scale, both in the range [0.001, 1000] in logarithmic steps of 10.

Platt's method (Platt, 1999) was used to generate class affiliation probabilities from the trained classifier. These continuously varying probabilities were then used to discriminate between responses and misses using both the Receiver Operator Characteristic (ROC) area under the curve (AUC) (Davis and Goadrich, 2006) and the classification accuracy as the percentage of correct predictions out of the total number of predictions. The classification accuracy was also computed by setting the class discrimination threshold as the optimal operating point of the ROC curve and calculating the percentage of correct predictions, using the threshold as a boundary between the two target classes. We used Wilcoxon signed rank tests (Gibbons and Chakraborti, 2011) to probe for significant differences between classification performances.

3.2.11. SPECTRAL POWER AND CONNECTIVITY ANALYSES

Spectral power and connectivity during responsiveness and unresponsiveness were investigated in both microstate-blind (pooling all samples, regardless of the labelled microstate) and microstate-wise (grouping samples by microstate label) analyses. Before microstate-wise segmentation, the power

spectral density was computed at each EEG sample between 1 and 20 Hz as the absolute value of the Hilbert transform (Marple, 1999) of the bandpass filtered data within windows of 0.25 Hz. We performed most of the analysis on 1 to 20 Hz and focused on theta and alpha power, whose ratio has been shown to track the onset of sleep (Šušmáková and Krakovská, 2007) and has been employed in other studies of drowsiness (Bareham et al., 2014) or impaired consciousness (Lechinger et al., 2013). For each channel in each recording, the spectral power at each frequency bin was averaged and normalised by the sum of spectral power within 1 to 20 Hz, thereby obtaining percentages of power contribution at every channel.

The connectivity within each pair of channels was analysed using the Weighted Phase Lag Index (WPLI) (Vinck et al., 2011), a connectivity measure based on the distribution of phase differences between signals. This measure is designed to correct for volume conduction by weighting the sign of the imaginary part of the cross-spectrum by the magnitude of the imaginary part itself, which is largest at 90° (where there is a delay between the signals) and minimal at 0°/180° (no delay, implying volume conduction). The WPLI been previously used to investigate brain connectivity during loss of consciousness (Chennu et al., 2016a, 2014; H. Lee et al., 2013). The WPLI was obtained by pooling over the Hilbert phase of each sample labelled as belonging to a particular microstate.

For both spectral power and connectivity, the median across channels was computed to obtain one value per microstate and frequency of interest.

To further assess topographical changes in connectivity, two sets representing anterior (AFz, Fz, FCz, AF7, AF3, F1, FC1, F3, FC3, F5, F7, AF8, AF4, F2, FC2, F4, FC4, F6, F8) and posterior (CPz, Pz, POz, Oz, P1, P2, PO3, PO4, O1, O2, P3, P5, P7, P4, P6, P8, CP3, CP1, CP2, CP4) electrodes were selected for analysis. Median WPLI connectivity was computed within the anterior and posterior groups separately for each participant.

3.3. RESULTS

3.3.1. BEHAVIOURAL FINDINGS

The distribution of responsiveness and reaction times over time confirmed that all the subjects were responsive for a minimum of six minutes in the beginning of the experimental session and became unresponsive at a later point. During the unresponsiveness period, participants predominantly reached sleep stage N1, and rarely N2, as detailed in (Kouider et al., 2014). Figure 3.1 shows the response reaction times and the misses in each participant, in addition to the selection of data for the subsequent microstate analysis. During responsive periods, most subjects had no more than one miss,

with a mean of 2.125% of all responses during this period being misses. The grand average of reaction times during the responsive period was 1.5s (S.D. = 0.7).

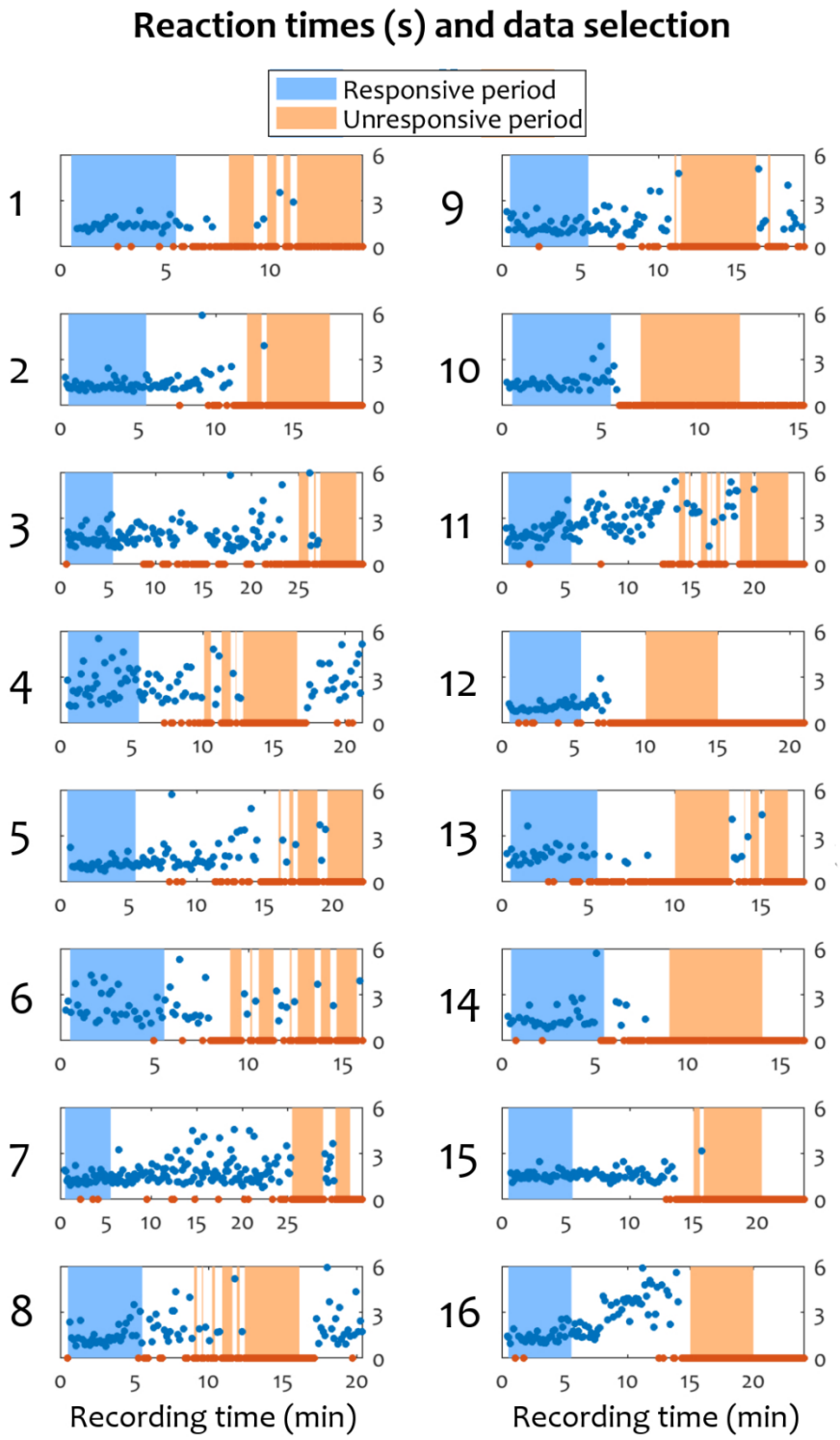

Figure 3.1 ***Reaction times and data segmentation into responsiveness and unresponsiveness for individual participants.*** *The horizontal axis represents recording time and the vertical axis represents reaction time in seconds. Blue markers indicate responses, while orange markers indicate misses. The blue area corresponds to the five-minute period of responsiveness, while the orange area corresponds to the five-minute period of unresponsiveness.*

3.3.2. SPECTRAL POWER AND CONNECTIVITY DYNAMICS

Before delving into microstate analyses, we characterised the spectral power and connectivity patterns during responsive and unresponsive periods. We performed a microstate-blind analysis focusing on previously reported changes related to early sleep, but also anaesthesia and disorders of consciousness, including the alteration of posterior, frontal and frontoparietal connectivity within and between frontal and parietal electrodes. We focused on alpha and theta frequencies, as the theta-alpha ratio has been shown to be a very good discriminator between wake and sleep stage 1 (Šušmáková and Krakovská, 2007), however we also confirmed that there were no significant differences in the means of power and median connectivity in beta (12-30 Hz) or gamma (30-40 Hz) between the responsive and unresponsive periods.

Based on the peaks present in alpha and theta bands in our data at 5.5 and 10 Hz (also see Figure 3.9 below) and in keeping with canonical definitions of EEG frequency bands, we defined the spectral frequencies of interest in alpha range at 9.5 to 10.5 Hz and the theta frequencies of interest at 5 to 6 Hz, for both power contributions and connectivity. Similarly narrow cut-offs have been defined in other studies involving a transition to sleep (Bareham et al., 2014) in order to avoid the smearing of the spectral peak.

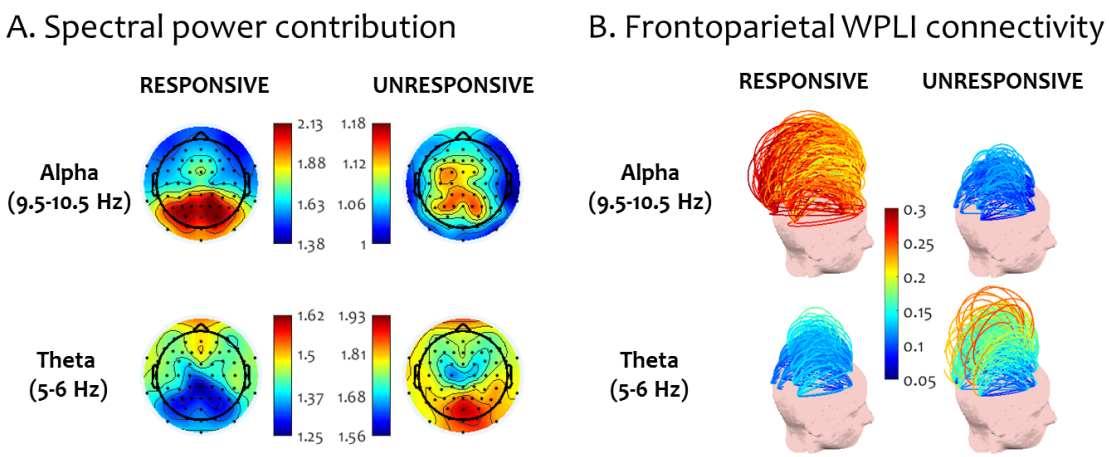

Figure 3.2 ***Spectral power topography and WPLI frontoparietal connectivity at alpha (9.5-10.5 Hz) and theta (5-6 Hz) peaks before and after the loss of responsiveness.*** *Values are averaged across participants.*

We observed a decrease in mean alpha power contribution ($t(1,15) = 3.34$, $p = 0.0044$, Cohen's $d = 0.83$) and an increase in mean theta power contribution ($t(1,15) = 7.1$, $p = 3.5e^{-6}$, Cohen's $d = 1.77$) going from responsiveness to unresponsiveness. As shown in Figure 3.3A, we noted an alpha peak in spectral power present around 10 Hz in the large majority of the participants during the responsive

period, which faded during the unresponsive period. Lower-frequency power in the theta frequency range increased during unresponsiveness.

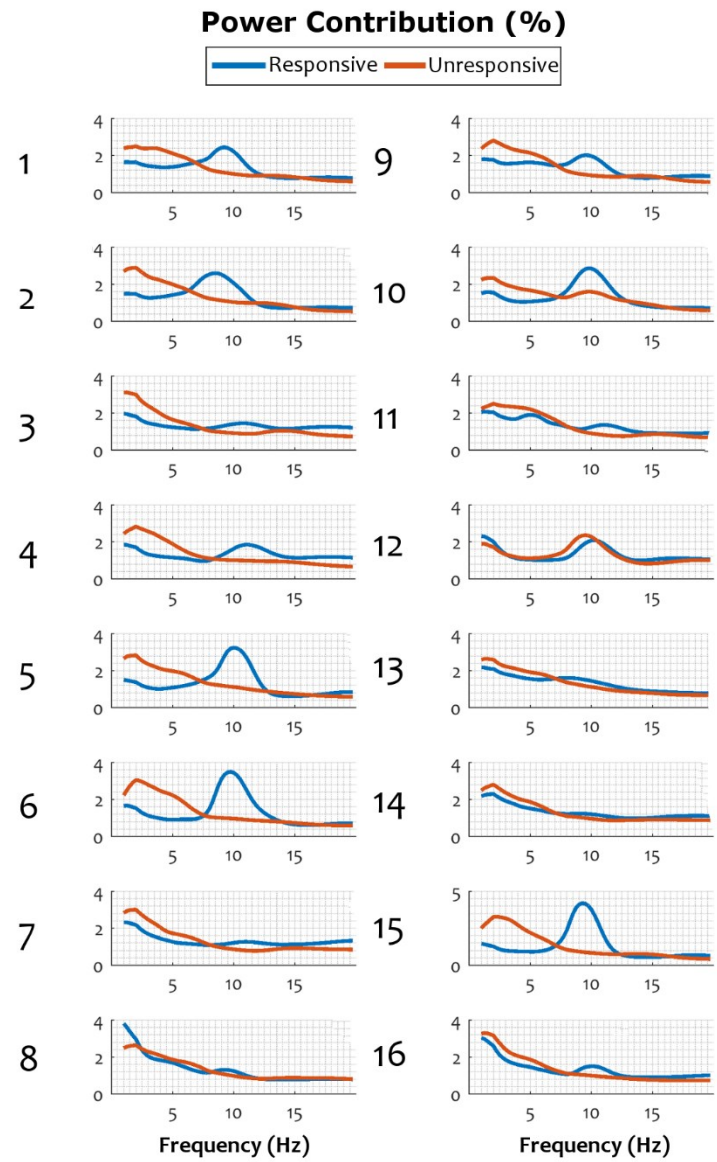

Figure 3.3 ***Individual subject spectral power contributions before and after loss of responsiveness.*** *For each subject, values are averaged over posterior channels.*

A single notable exception was Subject 12, whose alpha peak did not shift into theta range during the unresponsive period. This subject was not excluded from the analysis, as there was no evidence that the experiment instructions were not followed.

A grand average topographic plot of power at alpha and theta frequencies (Figure 3.2A) revealed that the highest alpha power was located in the posterior area during responsiveness. During unresponsiveness, theta power was highest in posterior channels.

Investigating frontoparietal connectivity in alpha and theta frequencies (Figure 3.2B) using the WPLI, we observed the disintegration of long-range alpha band connections between frontal and parietal electrodes going from responsiveness to unresponsiveness. A paired t-test confirmed that the median alpha connectivity between the anterior and posterior channels was significantly higher during responsiveness ($t(1, 15) = 3.4$, $p = 0.003$, Cohen's $d = 0.85$). At the same time, an overall increase in median frontoparietal connectivity was observed in theta frequencies during unresponsiveness, but this was not significant ($t(1, 15) = 0.4$, $p = 0.69$, Cohen's $d = 0.1$). The change in median connectivity is visible in most subjects, as shown in Figure 3.4.

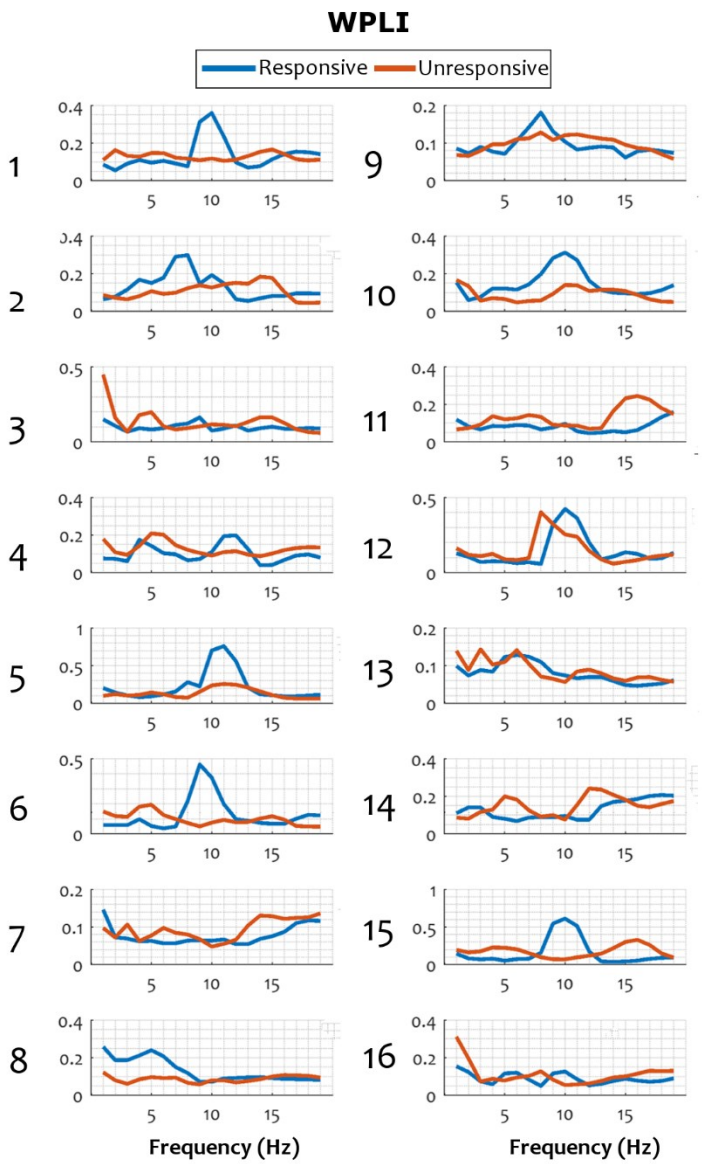

Figure 3.4 ***Median WPLI before and after loss of responsiveness due to drowsiness in individual subjects.***

3.3.3. MICROSTATE TOPOGRAPHIES

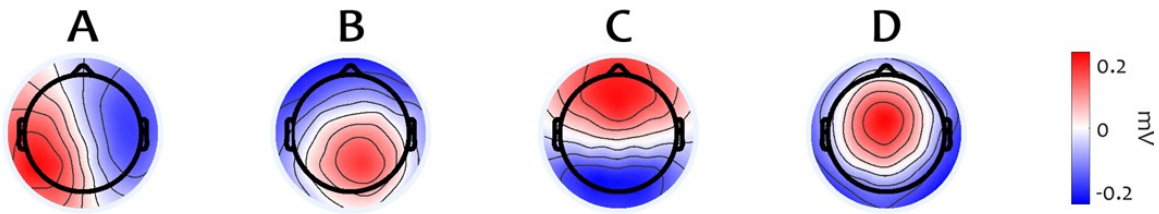

Figure 3.5 ***Microstate topographies computed across all subjects.***

It has previously been shown that microstate topographies are highly similar in wakefulness and sleep (Brodbeck et al., 2012). Hence, we applied the microstate clustering algorithm on the set of combined samples from the responsive and unresponsive periods from each subject, in order to obtain four microstate topographies. The resulting maps matched the four canonical microstate topographies commonly described in literature, denoted by letters A to D (Koenig et al., 2002) (Figure 3.5). A breakdown of microstate topographies obtained for individual participants is also shown in Figure 3.6.

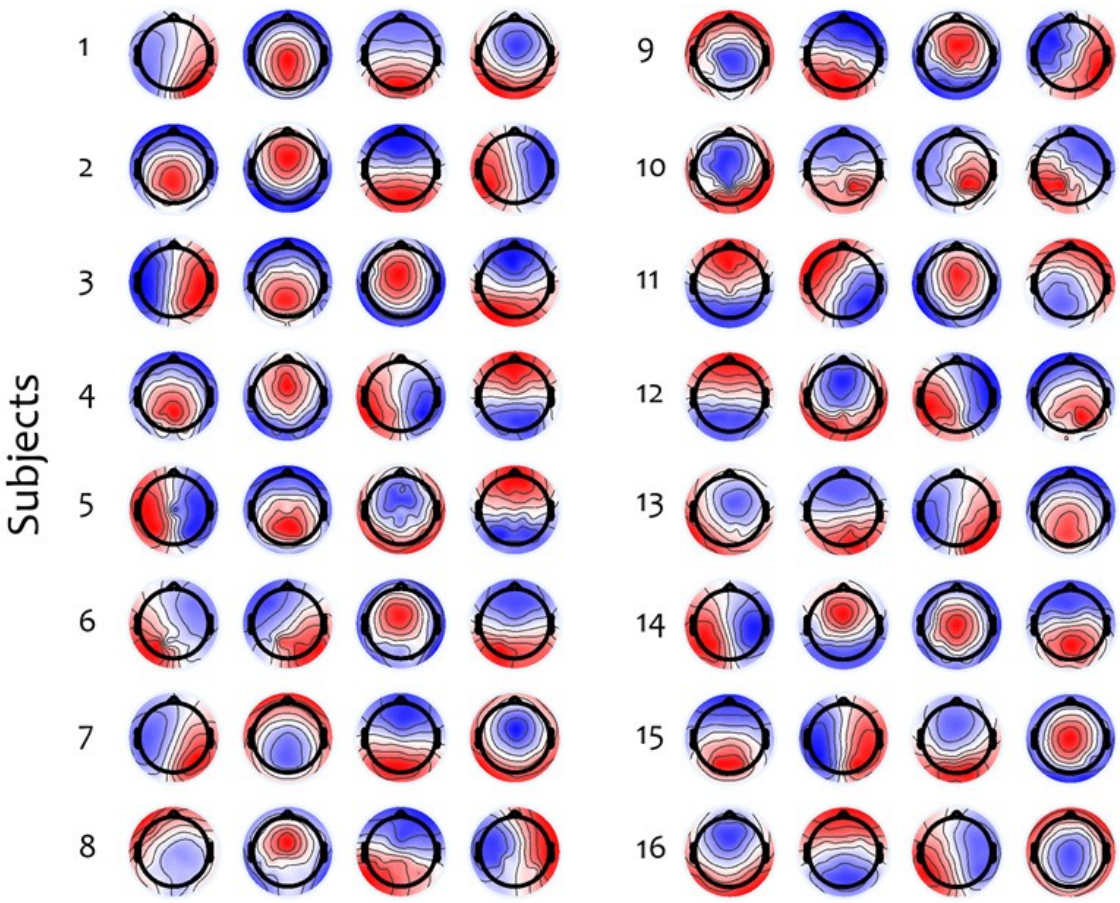

Figure 3.6 ***Microstate topographies in each subject, computed over the responsive and unresponsive periods.***

3.3.4. MICROSTATE PARAMETERS

Having established the topography of the canonical microstates, we next investigated whether the dynamics of the rapid succession of microstates in the EEG remains the same before and after the loss of responsiveness. We computed the duration, the temporal coverage and the global explained variance (GEV) of each microstate during responsiveness and during unresponsiveness (Figure 3.7).

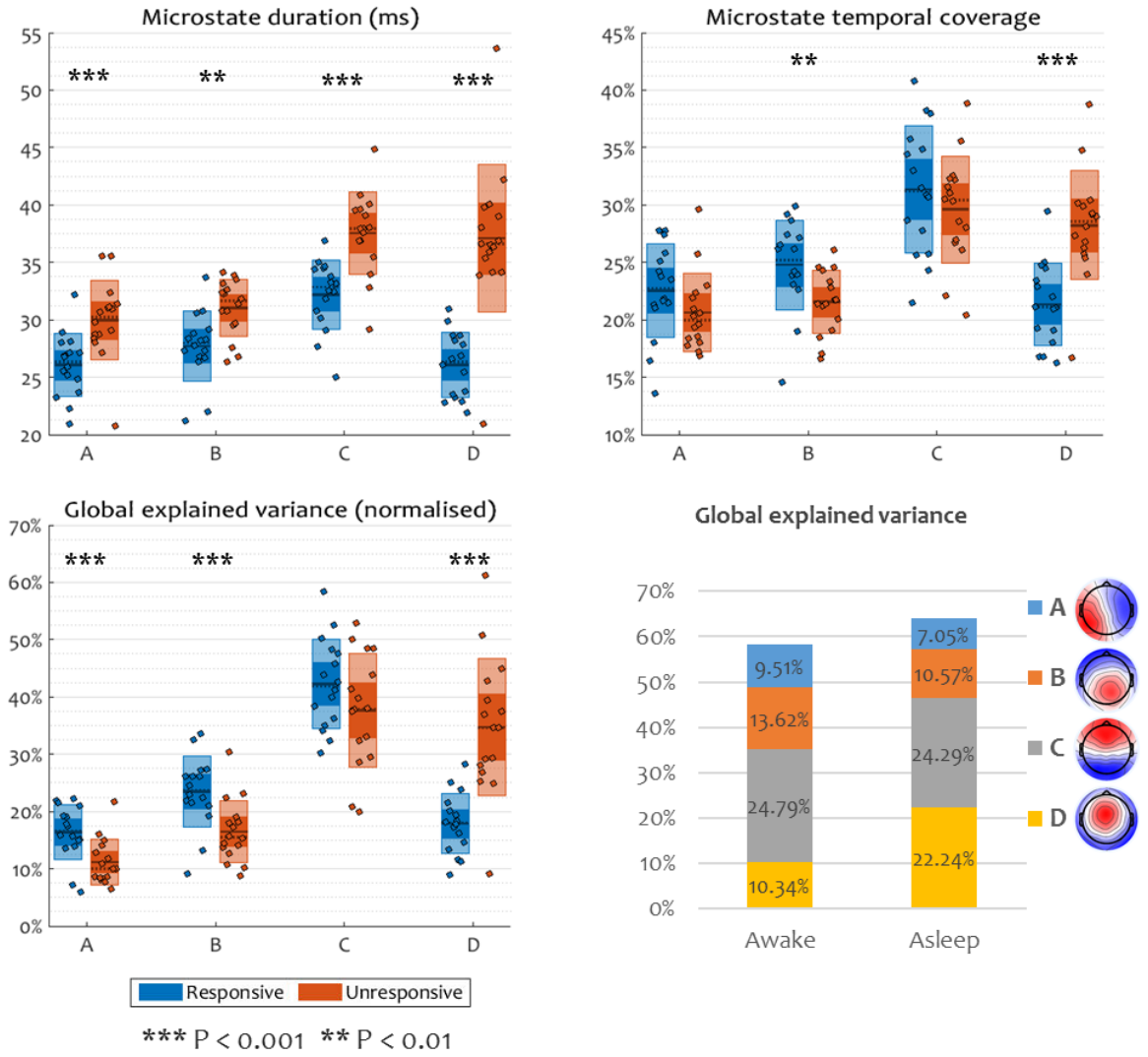

Figure 3.7 ***Microstate parameters before and after the loss of responsiveness in drowsiness.*** *Within each group, inner boxes represent the standard error of the mean, outer boxes represent the standard deviation, the mean is shown by a continuous line, the median is shown by a dotted line, and individual participant values are shown as dots. Asterisks show a significant main effect of state within a microstate.*

A repeated measures ANOVA with the microstate and the behavioural state (responsiveness and unresponsiveness) as factors found significant interactions between microstate and behavioural state

in all of the three microstate parameters investigated: duration ($F_{interaction}$ = 16.73, $P_{interaction}$ = $2e^{-7}$, Cohen's d = 2.11), temporal coverage ($F_{interaction}$ = 13.08, $P_{interaction}$ = $3e^{-6}$, Cohen's d = 1.86) and GEV ($F_{interaction}$ = 17.95, $P_{interaction}$ = $8e^{-8}$, Cohen's d = 2.18). Further exploring the simple effect of state on the parameters within each microstate, the ANOVA revealed that the duration of all microstates was significantly increased during unresponsiveness ($P_{state, A}$ = 0.0001, $P_{state, B}$ = 0.003, $P_{state, C}$ = 0.0001, $P_{state, D}$ = $3e^{-6}$), in agreement with previous literature (Brodbeck et al., 2012). Notably, microstate D had a striking increase in duration. At the same time, the temporal coverage of class D was significantly higher during unresponsiveness, whereas the coverage of microstate B was significantly lower during the same period ($P_{state, A}$ = 0.056, $P_{state, B}$ = 0.001, $P_{state, C}$ = 0.26, $P_{state, D}$ = $1e^{-5}$). Similarly, the GEV of microstate D was increased during unresponsiveness, while the GEV of microstates A and B were decreased ($P_{state, A}$ = 0.0002, $P_{state, B}$ = 0.0002, $P_{state, C}$ = 0.17, $P_{state, D}$ = $2e^{-5}$).

3.3.5. SINGLE-TRIAL RESPONSIVENESS PREDICTION

Having characterised the temporal changes in microstate dynamics before and after the loss of responsiveness, we proceeded to verify whether microstate parameters in the pre-stimulus window are able to dissociate responsiveness from unresponsiveness at individual trial level during the full recordings, and whether these properties could be generalised across subjects.

Out of all trials, 8% contained a button press event during the five seconds preceding each stimulus and were excluded from further analysis. The remaining data had a balanced distribution of 1078 responses and 1117 misses out of a total of 2195 trials.

Training a radial basis function kernel support-vector machine repeatedly on the combined-microstate and microstate-wise features to predict the binary outcome of a trial, as a response or a miss, using one-subject-out cross-validation, confirmed that microstate dynamics were able to predict responsiveness at individual trial level and across subjects, with a performance similar to that of the established theta-alpha ratio of spectral power (Figure 3.8).

Combining the duration, temporal coverage, and GEV of each microstate to obtain a 4 x 5 input feature vector or each trial achieved a mean AUC of 0.8552 (mean classification accuracy of 75.2%). In comparison, the theta-alpha ratio achieved a mean AUC of 0.8519 (mean classification accuracy of 74.24%). A Wilcoxon signed rank test did not find significant differences between these performance distributions. When combined, the microstate features and the theta-alpha ratio obtained a mean AUC 0.8622 (mean classification accuracy of 77.1%).

When used individually as input features for the classification, mean microstate duration performed remarkably well, achieving a mean AUC 0.8484 (mean classification accuracy of 76.1%). According to

Wilcoxon test, this was not significantly different from the classification performance of the combined microstate parameters. The duration of microstate D was significantly better at predicting responsiveness than microstates A-C ($p_{D\text{-}\{A,B,C\}}$={0.0005, 0.0006, 0.002).

It is worth noting that the one subject for whom the prediction performance was lower in the group was Subject 12, who was also the only participant whose alpha peak remained nearly unshifted after the loss of responsiveness (Figure 3.3, Figure 3.4).

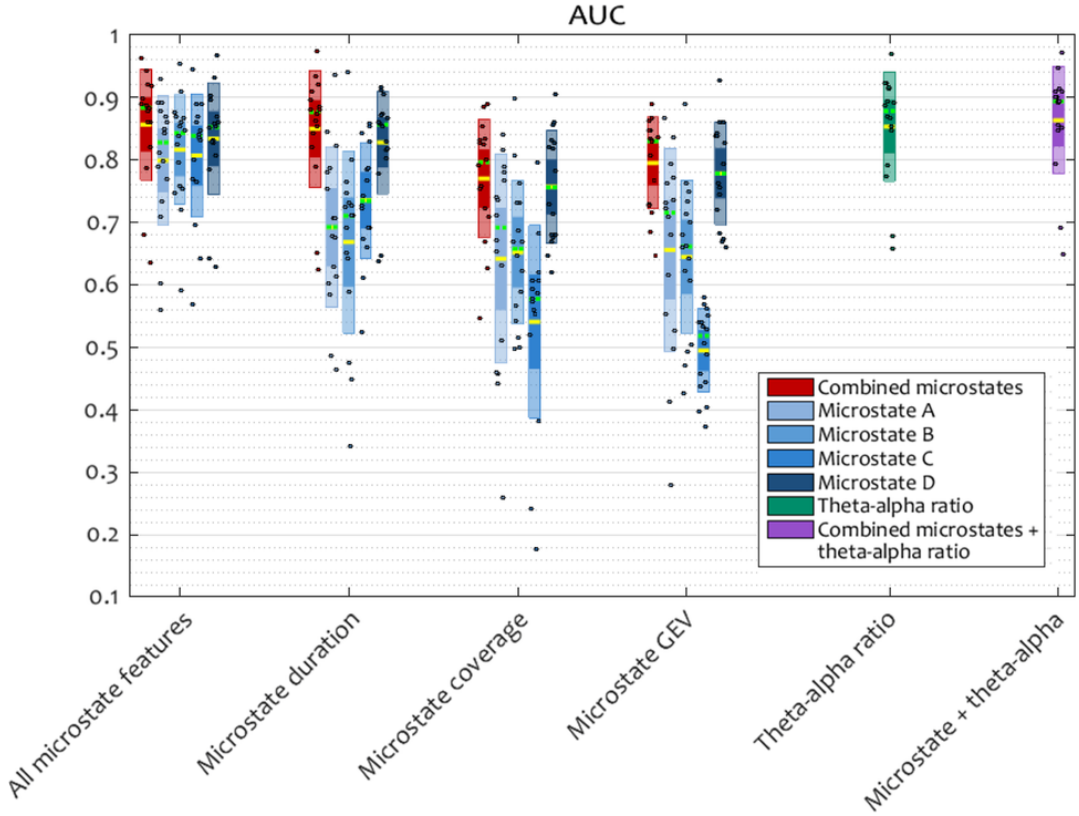

Figure 3.8 ***Classification performance, computed as the area under the ROC curve, for a support-vector machine (SVM) trained using 5 seconds of pre-stimulus data to classify responses and misses.*** *Input features are microstate parameters or the theta-alpha ratio, individually or combined. Within each group, inner boxes represent the standard error of the mean, outer boxes represent the standard deviation, the mean is shown by a yellow line, the median is shown by a green line, and individual participant values are shown as dots.*

Taken together, these results indicate that spatiotemporal microstate parameters characterising the pre-stimulus period are indeed informative of the ability of a subject to make a response, similar to the established theta-alpha ratio of the power spectral density. Confirming the initial findings of a more prominent presence of microstate D before the loss of responsiveness due to drowsiness, this microstate also appears to be particularly informative of the capacity of a subject to react to a stimulus. Crucially, these results are generalizable across subjects and valid at single trial level.

3.3.6. CONNECTIVITY DIFFERENCES BETWEEN MICROSTATES

Having established the characteristic temporal patterns exhibited by microstate sequences before and after drowsiness-induced loss of responsiveness, we next proceeded to investigate their relationship with the underlying spectral content of the EEG, and the modulation of this relationship as subjects become unresponsive. To this end, we investigated the power contributions and the WPLI connectivity computed across samples belonging to each microstate before and after the loss of responsiveness. While we do not assume a direct relation between neural sources of EEG microstates and EEG spectral power and connectivity, our aim is to assess whether the neural sources of microstates and sources of spectral measures covary at a fine temporal scale.

The spectral power contribution (Figure 3.9A) displayed the characteristic alpha peak around 10 Hz during the responsive period, which faded during the unresponsive period into high power at low frequencies. This pattern was similar during all microstates.

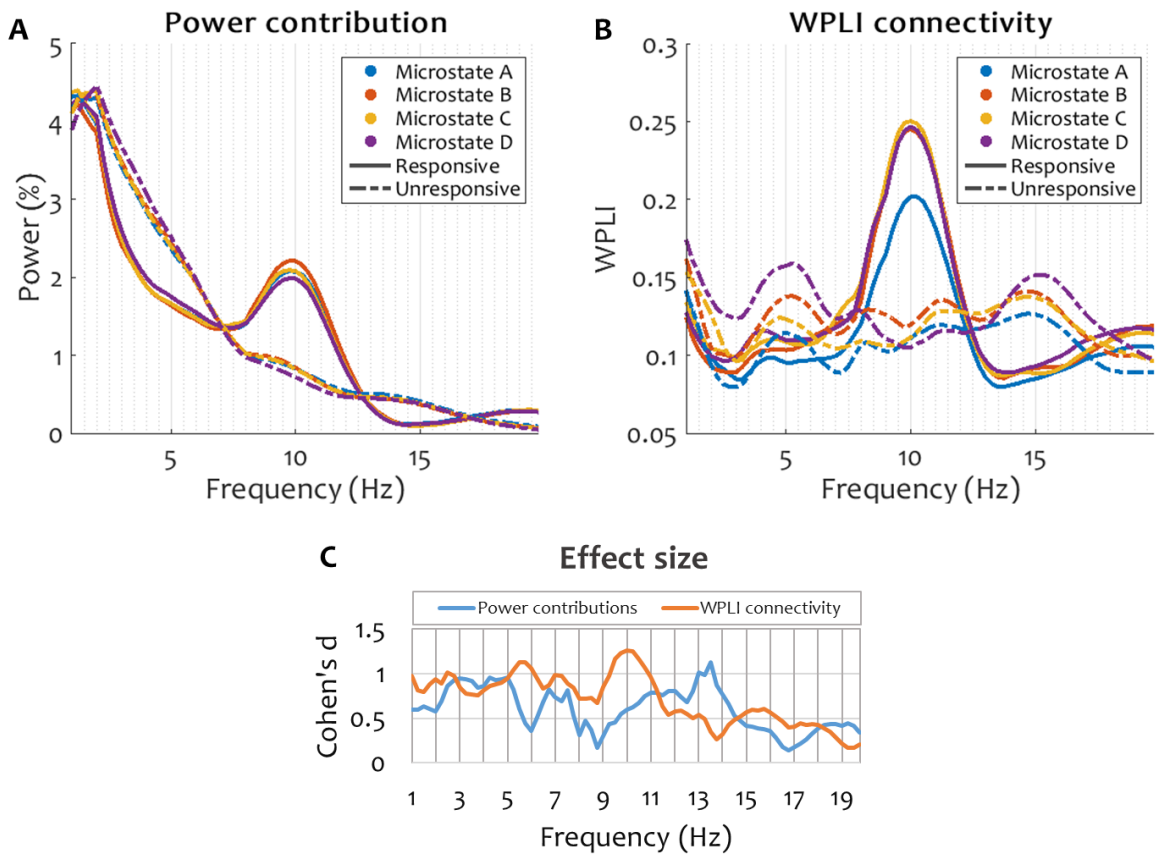

Figure 3.9. ***Spectral power contribution (panel A) and WPLI connectivity (panel B) captured during individual microstates before and after loss of responsiveness due to drowsiness.*** *Within each subject, for both power and connectivity, the median across channels was calculated. The figures show the grand average over all subjects. Panel C shows the main effect size, computed as Cohen's d, of the interaction between behavioural state and microstate at each frequency bin for power contributions and for connectivity.*

Likewise, spectral connectivity (Figure 3.9B) showed a peak at 10 Hz during responsiveness during all microstates, which faded during unresponsiveness. The only pattern dissociating between microstates during responsiveness was a decreased 10 Hz peak during microstate A. On the other hand, there was a noticeable difference in the level of connectivity during unresponsiveness between all microstate periods, with microstates D and A exhibiting the highest and the lowest connectivity, respectively.

The effect size of the interaction between microstate and behavioural state (responsiveness and unresponsiveness) computed individually at each frequency was indeed generally higher in connectivity than in power (Figure 3.9C). The effect size was largest in connectivity at 5.5 Hz and 10 Hz, corresponding to the theta and alpha peaks displayed during all microstates during the unresponsive and responsive periods, respectively. A peak in power contribution was also found at 13.5 Hz, potentially due to the emergence of sleep spindles at the onset of sleep.

We also attempted to use pre-stimulus WPLI connectivity levels at alpha and theta frequencies in order to train a classifier to predict responsiveness, using the same procedure as for the microstate spatiotemporal parameters. No classifiers could be obtained that exceeded a 60% mean accuracy, either microstate-wise or on the full set of pre-stimulus samples.

3.3.7. CONNECTIVITY DURING MICROSTATE D AFTER LOSS OF RESPONSIVENESS

Gathering from the evidence of increased temporal presence of microstate D after the loss of responsiveness, as well as the higher connectivity displayed during this microstate during unresponsiveness in comparison with the microstates A-C, we next sought to understand the spectral connectivity patterns captured during microstate D in the selected alpha and theta ranges during the unresponsiveness period.

Preliminary assessments of connectivity patterns during the four microstates during unresponsiveness revealed visual differences in anterior and posterior connectivity during microstate D as compared to microstates A-C. Considering previous literature (Morikawa et al., 1997; Tanaka et al., 2000, 1998; Wright et al., 1995) suggesting that key changes in connectivity related to the onset of sleep occur topographically in anterior and posterior scalp regions of interest (ROI), as well as frontoparietal networks having been proposed as a key signature of consciousness (Bor and Seth, 2012; Laureys and Schiff, 2012), we decided to investigate the within-anterior, within-posterior and between anterior-posterior connectivity during microstate D in comparison with microstates A-C. For this purpose, we performed three repeated measures ANOVA tests to compare the median connectivity during microstate D and that during each of the microstates A-C in each of the six conditions (two frequency bands X three scalp ROIs) during the unresponsive period. Within each condition, we corrected for the

false discovery rate across the three tests (D vs A, D vs B and D vs C) using Storey's procedure (Storey, 2002).

Figure 3.10 exemplifies the most prominent differences we found in connectivity between samples covered by microstate D and microstates A-C respectively, during unresponsiveness.

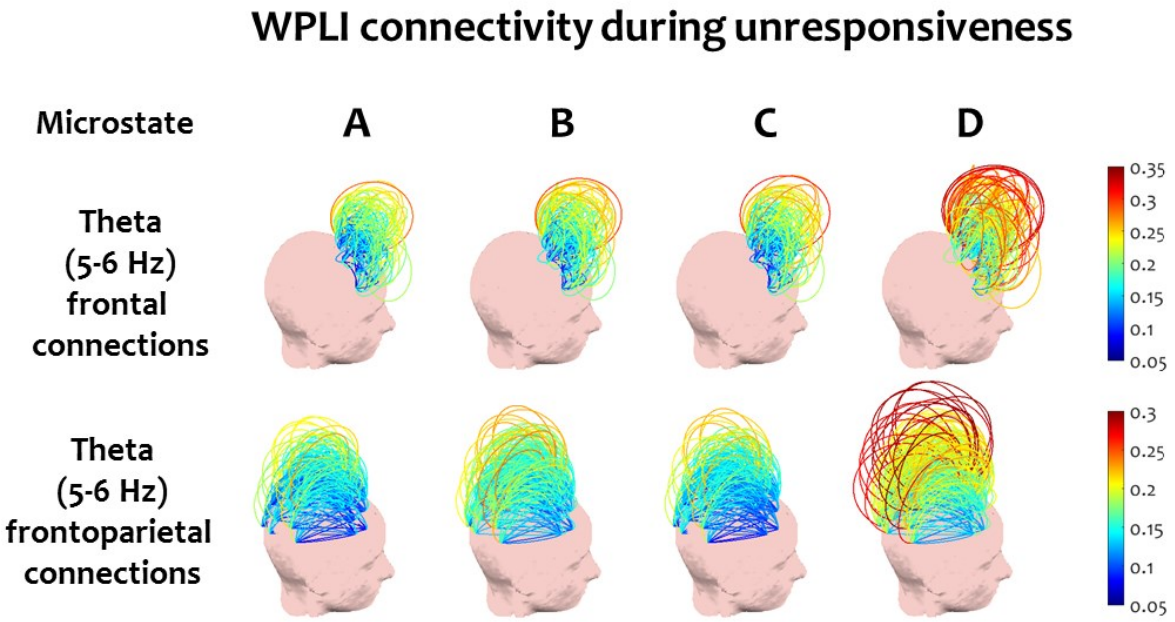

Figure 3.10 ***Frontal and frontoparietal WPLI connectivity at theta peak (5-6 Hz).*** *Microstate D captures significantly higher connectivity in these examples compared to microstates A-C.*

At the selected theta peak, the t-test results showed significantly higher median connectivity within the anterior region during microstate D compared to each of the other microstates ($P_{D\text{-}\{A,B,C\}}$ = {0.001, 0.008, 0.001}, $t_{D\text{-}\{A,B,C\}}$ = {3.958, 3.069, 4.088}, Cohen's $d_{D\text{-}\{A,B,C\}}$={0.990, 0.767, 1.022}). Median connectivity between the anterior and posterior regions was also significantly higher during microstate D than in microstates A and C ($P_{D\text{-}\{A,B,C\}}$ = {0.003, 0.297, 0.003}, $t_{D\text{-}\{A,B,C\}}$ = {3.578, 1.081, 3.392}, Cohen's $d_{D\text{-}\{A,B,C\}}$={0.894, 0.27, 0.848}). No significant differences were found in median connectivity within the posterior area.

Conversely, at the selected alpha peak, the repeated measures ANOVA showed significantly lower median connectivity within the posterior area during microstate D compared to microstates A-C ($P_{D\text{-}\{A,B,C\}}$ = {0.033, 0.037, 0.033}, $t_{D\text{-}\{A,B,C\}}$ = {2.686, 2.294, 2.559}, Cohen's $d_{D\text{-}\{A,B,C\}}$={0.672, 0.573, 0.67}). At the same time, microstate D captured significantly higher within-anterior median connectivity than microstate A ($P_{D\text{-}\{A,B,C\}}$ = {0.043, 0.617, 0.055}, $t_{D\text{-}\{A,B,C\}}$ = {2.769, 0.511, 2.297}, Cohen's $d_{D\text{-}\{A,B,C\}}$={0.692, 0.128, 0.574}). No significant difference in median connectivity between anterior and posterior regions was found during microstate D compared to microstates A-C.

These results confirmed that the timecourse of microstate D uniquely capture a simultaneous disintegration of posterior alpha connectivity and emergence of frontal theta connectivity, which is associated with the suppression of responsiveness at the onset of sleep.

3.4. DISCUSSION

3.4.1. SUMMARY

In this study, we used high-density EEG to explore the transient spatiotemporal and spectral dynamics of electrical brain activity before and after the loss of behavioural responsiveness due to drowsiness. Importantly, we examined the loss of responsiveness as participants became drowsy while performing a discrimination task. Hence, by design, our study is in contrast to and complements studies of resting-state brain activity in the absence of any task, which have often focused on investigating canonical sleep stages. Here, unresponsiveness – the failure to respond to the auditory cues elicited by increased drowsiness – provided an objective and non-invasive behavioural criterion in the transitional stage in between full wakefulness and early sleep.

We began by showing differences in spectral power and connectivity after the loss of responsiveness that have been previously shown to differentiate between healthy wakefulness and sleep, sedation and disorders of consciousness: a decrease in posterior alpha power and the emergence of theta power, as well as the disintegration of frontoparietal connectivity in alpha band. We then characterised the spatiotemporal parameters of the four canonical EEG microstates before and after the loss of responsiveness. We showed that microstate parameters not only correlate with behaviour at the group level, but also predict behaviour at the level of individual experimental trials. The ongoing microstate dynamics, particularly the properties of microstate D, before the onset of an auditory stimulus in an experimental trial significantly predicted the likelihood of a response to that auditory stimulus as participants transitioned towards sleep. Specifically, when microstate D occurred more often during the pre-stimulus period, participants were less likely to generate a response to the subsequent stimulus. This relationship highlights a possible functional role of this microstate in modulating behaviour, and the predictive power of this signature to define the capacity to consciously respond to abstract/semantic stimuli. Finally, we examined the spectral power and connectivity characteristics captured during the lifetimes of the four canonical EEG microstates. We discovered that while the distribution of spectral power remains the same across the temporal microstates, spectral connectivity has distinct profiles. We showed that this non-uniform pattern of connectivity across microstates is modulated specifically after the loss of responsiveness: the timecourse of microstate D captured significantly increased connectivity in the theta band after the loss of

responsiveness, underpinning a novel profile of interaction between the temporal sequence of microstates and spectral brain connectivity.

3.4.2. ALPHA POWER AND CONNECTIVITY CHARACTERISE RESPONSIVE WAKEFULNESS

Our analysis of EEG connectivity before microstate segmentation strengthens the evidence for the fundamental role of the frontoparietal alpha networks in sustaining a state of responsive wakefulness (Bor and Seth, 2012; Laureys and Schiff, 2012). Alpha band frontoparietal connections have also been shown to disintegrate in disorders of consciousness (Chennu et al., 2014) and sedation (Chennu et al., 2016a). Importantly, it is not the full disappearance of all frontoparietal connectivity that drives the loss of responsiveness, but specifically connectivity at alpha frequency. Indeed, literature shows that connectivity shifts from alpha into lower-frequency theta and delta frequencies as consciousness fades (Chennu et al., 2016a, 2014; Ogilvie, 2001; Tanaka et al., 2000, 1998; Wright et al., 1995). In the larger picture of states and levels of consciousness, our findings confirm long-range alpha networks as a common marker of consciousness, whether this impairment is natural (sleep), pathological (disorders of consciousness) or pharmacological (sedation).

3.4.3. MICROSTATE D PARAMETERS PREDICT RESPONSIVENESS

Upon examining the spatiotemporal parameters of the canonical EEG microstates, we found an increase in temporal coverage after the loss of responsiveness uniquely specific to microstate D, along with an increase in its global explained variance, as compared to responsive periods. While the duration of all microstates was longer during unresponsiveness, the duration of microstate D had a prominent relative increase. In contrast, the temporal coverage of microstate B decreased in the unresponsive period, as did the global explained variance of microstates A and B. Further, we demonstrated that the general state of awareness, as reflected in the ongoing dynamics of pre-stimulus EEG microstates, are indeed informative of the capacity of a subject to respond to a stimulus during drowsiness at an individual trial level. Again, the special significance of microstate D during unresponsiveness was visible from its increased ability to predict the likelihood of a response, in comparison with microstates A-C. In addition, we showed that the increase in duration of this microstate is the best predictor of responsiveness among all the microstate parameters.

Our usage of machine learning allows us to quantify the performance of the model using its discrimination accuracy, which speaks for the real-world applicability of the method (Breiman, 2001). Moreover, one-subject-out cross-validation allows us to infer that these results are generalizable across people. At the same time, as expected, individual variability caps the maximum possible

accuracy when predicting responsiveness. Our results suggest that this cap is around an accuracy of 75% (mean AUC around 0.85). Interestingly, the theta-alpha ratio, which we used as a baseline given its sensitivity as a sleep index (Šušmáková and Krakovská, 2007), achieved a similar classification accuracy as the microstate-based input features. Intriguingly, we were not able to use frontoparietal connectivity as a feature to train a suitable classifier for responsiveness during drowsiness, either considering or ignoring the microstate sequence, despite strong evidence of major connectivity changes occurring before and after the loss of responsiveness. This suggests that connectivity better predicts the level of consciousness estimated over longer time scales, whereas spatiotemporal microstate dynamics capture short-term changes in brain state that predict responsiveness.

3.4.4. MICROSTATE D CAPTURES A DISTINCT CONNECTIVITY PROFILE DURING UNRESPONSIVENESS

Alongside the distinctive increase in temporal coverage and duration of microstate D, we found a singular spectral connectivity pattern during this microstate after loss of responsiveness, indicating increased median connectivity in theta band, particularly in connections within frontal and between frontal and parietal electrodes. At the same time, median posterior connectivity during microstate D was reduced during unresponsiveness. Hence, the timecourse of microstate D appears to uniquely capture a connectivity pattern specific to deeper stages of sleep, in comparison with other microstates present during the same sleep stage. (Britz et al., 2010) have previously reported the lack of any interaction between temporal microstates of the brain and the spectral power of its oscillations, i.e, the spectral power profiles of EEG microstates do not differ from each other, a finding which we replicated. In contrast, we have shown that spectral connectivity presents a significant interaction with temporal microstate dynamics, underpinned by the connectivity captured by microstate D.

There currently exists no consensus on the meaning of individual microstates in terms of their neural generators. However, microstate D has occasionally been linked to attentional networks. In a study of fMRI resting-state networks, (Britz et al., 2010) showed a higher correlation of microstate D with ventral and dorsal frontal-parietal networks, functionally associated with attention switching and directing attention towards external salient stimuli. A decreased duration of this microstate has been reported in schizophrenia (Koenig et al., 1999; Lehmann et al., 2005; Nishida et al., 2013; Tomescu et al., 2014) and hallucination (Kindler et al., 2011) – two conditions involving impairments in task switching and attention (Collerton et al., 2005; Cornblatt and Keilp, 1994). An investigation of modalities of thinking found an increased microstate D duration in resting-state compared to visual and verbal task periods (Milz et al., 2015); this was also interpreted as a confirmation of the previously-mentioned study by (Britz et al., 2010) due to a higher probability of attention switching during rest

(high microstate D duration), as opposed to performing a single goal-oriented task (lower microstate D duration). On the other hand, (Seitzman et al., 2016) have found an increased duration of microstate D during a cognitive task as compared to wakeful rest.

Given the weak evidence in the literature associating microstate D with task-related attention networks, we are cautious in interpreting our findings on this basis. A previous study on the same data (Kouider et al., 2014) found that a correct response to stimuli is still prepared during unresponsiveness, suggesting preserved attention. It is possible that our findings indicate more demand from attention networks as drowsiness increases and subjects become unable to respond to the task. In study of microstates during sleep in the absence of any task, (Brodbeck et al., 2012) did not observe an increase in this microstate during sleep. This suggests that microstate D might indeed be specifically related to the experimental task. Further, this interpretation is compatible with a study by Katayama et al. (Katayama et al., 2007), which found that the duration of microstate D was increased in light (but not deep) hypnosis, a state which produces similar EEG patterns to sleep-induced unresponsiveness (Barker and Burgwin, 1949).

Nonetheless, the spatiotemporal and spectral connectivity dynamics observed in microstate D after the loss of responsiveness yield an important insight into the dynamics of the transition to sleep. While connectivity averaged during all microstates reflects typical changes commonly found in the loss of consciousness in the onset of sleep, anaesthesia or disorders of consciousness – weaker alpha and stronger theta long-range networks – the individual timecourse of microstate D captures this change in connectivity to a significantly larger extent than microstates A-C. In other words, during microstate D, alpha connectivity is weaker, while frontoparietal theta connectivity is stronger. This happens despite microstate D having a duration no longer than 40ms. This suggests that, after the loss of responsiveness, the process of falling asleep is not necessarily linear, but rather consists of an interplay between distinct networks, captured by different microstates, which are at different points along the transition between wakeful and asleep modes of operation. This finding might lend itself to explaining one of the current riddles of sleep research: why is it that, despite the establishment of a series of clear EEG markers delimiting wake and several stages of sleep, finding an EEG-based threshold to separate between the subjective intuition of being awake or asleep has not yet been achieved? Indeed, it has been reported by Hori et al. (1994) that 26% of all subjects stated that they had been awake at times when their EEG was classified as stage 2 sleep, which is commonly used to define 'true sleep' (Ogilvie, 2001). The rapid fluctuation of brain networks, some of which are closer to wakefulness (during microstates A-C) and others closer to sleep (during microstate D) could be the reason why our momentary introspective state of being 'awake' and 'asleep' might not concur with a coarse-grained assessment of EEG over many seconds of data, as usually done during the identification of sleep

stages. Instead, our findings here highlight that further research should focus on the rapidly changing dynamics of brain networks that appear to capture key dynamics relevant to our behavioural and perhaps even introspective state, as we drift into unconsciousness.

3.5. ACKNOWLEDGEMENTS

We thank Louise Goupil for collecting the data for this experiment.

CHAPTER 4

SEDATION

The previous chapter explored spectral and topographical markers of the loss of responsiveness during the transition to unconsciousness while falling asleep. The onset of sleep is a natural process where consciousness is lost and regained in response to environmental and internal cues. But how is consciousness lost as a consequence of an external pharmacological agent? This chapter investigates this question in a group of healthy adults who were administered sedative doses of propofol. As a proxy to assess their level of consciousness, responsiveness to simple auditory stimuli is employed once again. On the same dataset, a previous study (Chennu et al., 2016a) found that spectral power and connectivity showed highly similar changes to those found during the onset of sleep in the previous chapter: a decrease in alpha connectivity and power, accompanied by a corresponding increase at lower frequencies. This chapter focuses on predictions originating in recent theories of consciousness and investigates information-theoretical measures of integration and differentiation of EEG signals in order to find signatures of drug level and responsiveness. Previous findings from other studies (Schartner et al., 2015) regarding the decrease of EEG signal diversity with unconsciousness induced by sedation are confirmed. Further, it is discovered that drug level and responsiveness are best explained by distinct neural signatures combining differentiation and integration, adding to evidence of a similar dissociation previously found using fMRI (Barttfeld et al., 2015).

4.1. INTRODUCTION

Recent theories that seek to uncover the neurodynamics of consciousness have gained practical interest in both theoretical and clinical neuroscience (Dehaene et al., 2014; Koch et al., 2016; Tsuchiya, 2017). Several such theories, such as the integration information theory (IIT) (Oizumi et al., 2014; Tononi, 2004), the dynamic core hypothesis (Tononi and Edelman, 1998) or the metastable brain theory (Tognoli and Kelso, 2014), stipulate that full consciousness requires an optimal balance of information integration and differentiation within the neural system. Due to limitations of theoretical and computational nature, no exact measures directly derived from theory can currently be computed for the whole human brain (Oizumi et al., 2014). However, practical measures inspired by these theories of consciousness have been developed to explore the dynamics of neural activity

differentiation and integration in healthy and impaired states of consciousness (Barrett and Seth, 2011; Seth et al., 2011).

Pharmacologically-induced loss of consciousness is one condition where recent theories of consciousness can provide a valuable framework (MacDonald et al., 2015; Marchant et al., 2014). Understanding the micro- and macroscopic level changes in neural activity that underlie anaesthesia (the complete loss of consciousness) and sedation (an impaired state of consciousness) is still a challenge for neuroscience, for both theoretical and clinical purposes. Current electrophysiological markers of anaesthesia depth, such as the bispectral index, do not always prevent intraoperative awareness (Russell, 2013), with 0.13% cases of accidental awareness estimated to occur (Sebel et al., 2004), although this varies with surgery type and anaesthetic (Pandit et al., 2014). The development of better methods to track anaesthesia is hence desirable.

The above theories predict that consciousness impairment is accompanied by decreased neural information differentiation and integration. This can help us design better measures for tracking anaesthesia and sedation. In line with this, it has been shown that the capacity of long-range interaction across the neural system decreases during sedation (Alkire et al., 2008; Boveroux et al., 2010b; Koch et al., 2016; Lee et al., 2009b; U. Lee et al., 2013; Mashour, 2004; Monti et al., 2013), although local activity may be preserved (Lewis et al., 2012), particularly in sensory networks (Bonhomme et al., 2012; Boveroux et al., 2010b). EEG studies have also suggested that anaesthesia reduces the complexity of the neural signal (Wang et al., 2017) by decreasing the repertoire of discriminable states available to the neural system (Hudetz et al., 2015; Lee et al., 2017). This idea is also supported by connectivity changes observed in the fMRI during sedation (Stamatakis et al., 2010). Overall, quantifying information differentiation and integration in neural activity seems to be a fruitful approach in elucidating the brain changes underlying sedation and anaesthesia (Alkire et al., 2008).

While anaesthesia has been studied extensively, there is currently limited work that addresses the neural changes occurring during the loss of responsiveness during sedation, which also marks the beginning of the transition to full unconsciousness. At similar sedative doses of anaesthetic, individuals respond differently, with some remaining fully awake and others losing responsiveness (Chennu et al., 2016a). Although unresponsiveness does not imply a complete lack of awareness (Boly et al., 2013a; Sanders et al., 2012), understanding the neural dynamics underlying the loss of responsiveness in the early stages of anaesthesia can give us important insight into fine-grained changes that underlie the loss of consciousness. In practice, the degree of behavioural responsiveness of the patient (Boly et al., 2013a) is widely used as a clinical marker of the level of consciousness (Laureys, 2005). Following the same behavioural criterium, here we investigate a level of sedation

where only a proportion of the subjects (37.5%) lose responsiveness. For this purpose, we employ two information-theoretical measures to quantify information differentiation and integration in neural dynamics.

Lempel-Ziv (LZ) complexity (Lempel and Ziv, 1976) is a promising information-theoretical method for assessing information differentiation in the EEG signal. LZ complexity quantifies the diversity of a binary string by counting the number of different patterns it contains. Interest in LZ complexity has been recently revived by the introduction of the perturbational complexity index (Casali et al., 2013), which can reliably distinguish between states of consciousness, including anaesthesia (Sarasso et al., 2015), by computing the complexity of the response elicited by TMS. LZ complexity has also been shown to track consciousness state when applied on spontaneous electrophysiological activity in anaesthesia (Bai et al., 2015; Hudetz et al., 2016; Schartner et al., 2015; Zhang et al., 2001), sleep (Abásolo et al., 2015; Andrillon et al., 2016; Schartner et al., 2017b), disorders of consciousness (Sitt et al., 2014) and psychedelic experience (Schartner et al., 2017a). Hence, it has proven to be a simple and efficient index for monitoring sedation and anaesthesia. Two different variants of LZ complexity have been described: one that quantifies local, channel-wise complexity (LZS/LZSUM) and one that quantifies the complexity of the global topography of the scalp over time (LZC). These two variants appear to occasionally track different aspects of consciousness (Schartner et al., 2017a).

The weighted Symbolic Mutual Information (wSMI) index (King et al., 2013) is a recently-introduced information-theoretical method for assessing information integration in the EEG. It measures information sharing between two signals by looking at their joint, nonlinear fluctuations, and corrects for volume conduction by discarding identical or opposite symbols. This measure has been shown to discriminate patients with disorders of consciousness (Claassen et al., 2016; King et al., 2013; Sitt et al., 2014) and therefore seems to be a promising information-theoretical measure for tracking consciousness levels. To our knowledge, this measure has not yet been applied to sedation and anaesthesia.

In this context, we employ these two information-theoretical measures on high-density EEG data in order to characterise the differentiation and integration of neural information during the administration of propofol at doses that impair motor responsiveness to a simple auditory perceptual discrimination task. Starting with an awake baseline, as sedation progressed to mild and then to moderate levels, the subjects became gradually less responsive during the infusion, allowing us to study two groups of subjects: those that stay responsive and those that become unresponsive during sedation. Alongside, the concentration of drug in blood plasma was measured at each level of sedation. We applied the LZ complexity and the wSMI on the EEG as measures of neural information

differentiation and integration. In line with previous findings, we hypothesised that neural information integration and differentiation would track the level of drug and the participants' responsiveness. We start by performing a group level analysis to find an interaction between responsiveness and sedation depth, and follow up with a more fine-grained regression analysis to find the best subset of predictors for drug concentration, response rate and median reaction times. Interestingly, we find a non-overlapping subset of predictors for drug level and responsiveness, with both predictors requiring a simultaneous combination of both complexity and integration measurement. We conclude that drug concentration and responsiveness have distinct neural correlates, with drug level best predicted by spatial complexity and high frequency mutual information sharing, whereas responsiveness is best predicted by temporal complexity and low-frequency mutual information sharing. Further, we find that the group that remains responsive during sedation displays increased anterior single-channel LZ complexity compared to the group that became unresponsive, whereas the latter shows a loss of complexity in the anterior area, emphasising the importance of frontal complexity in the preservation of responsiveness despite the influence of the sedative.

4.2. METHODS

4.2.1. EXPERIMENTAL PROCEDURE

A sample of 25 neurologically healthy subjects participated in the sedation experiment. Nine subjects were excluded due to incomplete or corrupted data. The remaining 16 subjects (9 female) had a mean age of 30.9 (S.D. = 10.9). All participants gave written consent prior to participating in the experiment. The experiment was approved by the Cambridge Psychology Research Ethics Committee.

Following a wakefulness baseline period lasting 25-30 minutes, a computerised syringe driver (Alaris Asena PK, Carefusion, Berkshire, UK) was used to induce sedation by specifying the target plasma level concentration of propofol. The target concentrations were 0.6 mg/L for mild sedation and 1.2 mg/L for moderate sedation. The aim of the mild sedation stage was to induce a relaxed but responsive state, whereas moderate sedation was aimed at a threshold where a proportion of the participants would become unresponsive. A recovery time of 20 minutes was allowed after the end of sedation, based on pharmacokinetic software simulations to estimate the time necessary to reach a plasma concentration level of zero.

At each of the target levels (baseline, mild sedation, moderate sedation and recovery), a resting-state period of approximately 7 minutes was recorded, followed by an auditory discrimination task. The participants performed a perceptual discrimination task where they were asked to classify the stimulus as a noise or a buzz by pressing a button with their left or right hand. From the presentation

of the stimulus, a window of 5 seconds was allowed for making a response. The lack of a response in this interval was considered a miss. The task consisted of 40 stimuli in total at each sedation level.

4.2.2. EEG DATA RECORDING

EEG data was recorded during the experiment from a high-density 128-channel net using a Net Amps 300 amplifier (ElectricalGeodesics Inc., Eugene, Oregon, USA). The data was referenced at the vertex and sampled at 250 Hz. Channels situated on the cheeks, forehead and neck were excluded in order to minimise muscle-related artefacts. A total of 90 channels over the scalp were retained for further analysis.

4.2.3. EEG PRE-PROCESSING

EEG data processing was performed in MATLAB (The MathWorks, Inc., Natick, Massachusetts, US) with custom scripts using functions provided by the EEGLAB toolbox (Delorme and Makeig, 2004).

The resting-state sessions at baseline, mild sedation, moderate sedation and recovery were retained for EEG analysis, whereas the task periods were used to extract behavioural data. For each session, the data was filtered between 0.5 and 40 Hz. All sessions were concatenated in order and Independent Component Analysis (ICA) was performed. ICA components indicating muscular artefacts were removed. Finally, the data was visually inspected for channel artefacts and noisy channels were interpolated across all sessions.

4.2.4. MEASURING COMPLEXITY

To quantify the diversity of patterns present in the electric signal, an adapted version of the Lempel-Ziv-Welch (Welch, 1984) (LZW) algorithm was used (Box 4.1). The algorithm assigns a complexity number to a binary string by counting the number of unique patterns it contains. For this study, the algorithm, as developed by (Schartner et al., 2015), was implemented in C++ and used with custom MATLAB scripts. These scripts are available online at https://github.com/iulia-m-comsa/EEG.

The data was divided into 10-second epochs and baseline-corrected by subtracting individual channel means. This epoch size is similar to that used in other studies computing the LZ complexity on human EEG (Schartner et al., 2017b, 2015) and provides a good balance between smaller epochs that result in noisier values and larger epochs that result in less temporal accuracy. Epochs with variance higher than 200 uV were rejected. This threshold was established by visual inspection of the EEG data, with the aim to remove artefacts while also preserving as much data as possible. A mean (S.D.) of 39.4 (1.5), 39.5 (1.4), 37 (4.7), 39.5 (2.7) epochs were retained for the baseline, mild sedation, moderate sedation and recovery sessions respectively. The Hilbert transform was used to obtain the instantaneous

amplitude of the signal. In every subject, complexity was computed for every epoch and then averaged per state.

To apply the LZW algorithm on a data epoch, each channel was converted into a binary string using the channel mean as threshold. On the resulting binary matrix of channel values over time, we investigated two types of complexity, as explained below.

- *Spatial complexity (LZC)*. This measure was computed by concatenating the data in a spatial manner (observation by observation), with an observation consisting of all ordered channel values at a single time frame. The LZW algorithm was applied on the concatenated string. This measure is termed LZC in (Schartner et al., 2015). This measure represents the diversity over time in the spatial patterns of the EEG, hence representing a *global* measure of diversity.
- *Temporal complexity (LZS)*. This measure was computed by concatenating the data in a temporal manner (timeseries by timeseries), therefore preserving the temporal structure at each individual channel. This measure was called LZS in (Schartner et al., 2015). An alternative measure, called LZSUM in (Schartner et al., 2015), is obtained as the mean of the LZW complexity computed on single channels. The individual LZW value at each channel can also provide a topographical map of complexity. Temporal complexity measures the diversity of patterns found in individual channels, hence providing a measure of *local* diversity.

Lempel-Ziv-Welch (LZW) complexity algorithm

Input: a sequence S of binarised EEG data (k samples x c channels)

Output: a number representing the raw LZW complexity of the sequence

1. Flatten S column-by-column (for LZC) or row-by-row (for LZT)
2. Initialise the dictionary of patterns D = {{0} and {1}}
3. Initialise current_sequence = S[0]
4. Initialise i = 0
5. *while* i < length(S) *do*:
    6. Append S[i] to current_sequence.
    7. *if* current_sequence ∉ D *then* add it to D
    8. i = i + 1
9. Return length(D)

Box 4.1 ***Lempel-Ziv-Welch algorithm.*** *Returns the number of unique patterns found in the input data.*

To obtain a normalised complexity value for every epoch, two types of normalisations were used sequentially. The first method consists dividing the complexity of the given string by that of a shuffled version of the same string. The shuffled version represents the maximal diversity of a string composed of the same characters. To obtain a representative maximal complexity for normalisation, the original string was randomly shuffled 10 times and the mean complexity of the obtained strings was used to divide the complexity value of the original string. We consider 10 random shuffles to give an appropriate estimation of maximal complexity. It has been previously shown that even a single random shuffling is sufficient to approximate the complexity of a random sequence for 10-second epochs at the same sampling rate (Schartner et al., 2015). A string of maximal diversity would hence have a complexity value of 1.

The second normalisation seeks to ensure that any changes in complexity cannot be explained by changes in the power spectrum. For this purpose, phase-randomised surrogate data with the same Fourier spectrum as the original data is obtained as explained in (Theiler et al., 1992). The Fourier transform of the data is computed and the phase of each complex frequency component is randomised by multiplying it with a uniformly random phase between 0 and 2π, while the amplitude of the component is preserved. The inverse Fourier transform is then applied to obtain the surrogate data, on which complexity is computed. Thus, we obtain the maximal complexity of a string with the same spectral profile. As before, the mean complexity of 10 such randomisations were used for normalising the complexity of the original signal. This method aims to ensure that any changes in complexity that can be explained by the power spectrum only are removed from the result. Our measures hence correspond to the LZ_N measures in (Schartner et al., 2017a).

As an alternative and conservative measure to check that complexity is not influenced by the power spectrum, notch filters in theta, alpha and beta bands, which have been shown to be altered during propofol sedation, were applied in turn to the data before recomputing the LZ complexity and verifying the effects obtained in the original results.

4.2.5. MEASURING INFORMATION SHARING

In order to measure the dynamics of information integration during sedation, we calculated the weighted symbolic mutual information (wSMI) as introduced by (King et al., 2013). The wSMI assigns a proportional value representing the co-occurrence of similar, nontrivial patterns between pairs of channels, thereby providing a measure of information sharing. To ensure the measure captures a true reflection of underlying neural information exchange, signal patterns that might result from common sources are discarded.

The data was split into 1-second epochs and baseline-corrected by subtracting individual channel means. Epochs with variance higher than 200 uV were rejected. A mean (S.D.) of 393 (13.5), 396.2 (11.6), 371.6 (28.5), 390.7 (32.1) epochs were retained for the baseline, mild sedation, moderate sedation and recovery sessions respectively. The wSMI was computed for every pair of channels at every epoch. Due to the high variance of this measure (as shown in the Results), a trimmed mean wSMI per epoch was obtained by averaging across all pair values after excluding the 10% highest and 10% lowest values. This procedure has been used in other recent studies involving the wSMI (Engemann et al., 2018). However, we also confirmed the results were very similar and statistically equivalent if the median of the wSMI was used instead of the trimmed mean.

We investigated four temporal separation parameters (τ = 4, 8, 16 and 32 samples) corresponding to maximal frequencies of 20.8 Hz, 10.41 Hz, 5.2 Hz, and 2.6 Hz respectively. These upper limits correspond approximately to the beta, alpha, theta and delta frequency bands, all of which have been shown to capture electrophysiological changes induced by propofol sedation (Chennu et al., 2016a; Murphy et al., 2011; Purdon et al., 2013; Seifert et al., 1993).

4.2.6. GROUP DIFFERENCES ANALYSIS

For the initial analysis, participants were classified as responsive or unresponsive based on their ability to respond to the noise/buzz stimuli during moderate sedation compared to the baseline period. The number of hits and misses in the perceptual discrimination task was counted during each of the four sedation levels. A binomial distribution was fitted to each state and the 95% confidence intervals of each distribution were obtained. Each participant was classified as unresponsive if the confidence interval of responses was lower and non-overlapping with the confidence interval at baseline, and responsive otherwise.

A repeated measures ANOVA (Hogg and Ledolter, 1987) was used to assess the interaction between level of sedation (baseline, mild, moderate and recovery) and group (responsive and unresponsive) for each measure of interest. To test for differences between the two groups at the same sedation level, and between sedation levels within the same group, comparisons of marginal means were performed and corrected for multiple comparisons using the Tukey-Kramer method (Tukey, 1949). Where Mauchly's test of sphericity (Mauchly, 1940) was violated, the p-value was corrected using the Greenhouse-Geisser procedure (Greenhouse and Geisser, 1959). A conventional threshold of 0.05 was applied to the p-value to assess significance.

4.2.7. OPTIMAL PREDICTORS

A regression analysis was performed to assess the predictability of drug concentration, responsiveness and RT using signal complexity and information integration. Complexity and WSMI values at mild sedation, moderate sedation and recovery were normalised by the value at baseline. To minimise individual differences in drug uptake and increase sample independence at mild and moderate sedation, drug level was normalised by the value at recovery. Responsiveness was computed as the number of responses out of a maximum of 40. As we noted that subjects were still adjusting to the task during the baseline period (the number of misses was in some cases larger at baseline than in all other stages), responsiveness was also normalised by the value at recovery. Likewise, reaction time (RT) was computed as the mean of response times normalised by the mean response time at recovery level. Points classified as unresponsive using the procedure described in the previous section were excluded from RT analysis.

The performance of LZW (LZC, LZS, LZSUM) and wSMI (τ = 4, 8, 16, 32) in predicting drug level, responsiveness and RT was assessed using a generalised linear model. Drug concentration and RT were modelled as normal distributions. Responsiveness was modelled as a binomial distribution specifying the number of responses recorded out of the maximum of 40 and using the logit linking function. The measures were entered individually as predictors and the coefficient of determination (R^2) was used to compare the variance explained by each predictor.

To select the best predictors for each dependent variable, we performed a generalised linear regression with Lasso regularisation (Tibshirani, 2011). In Lasso regularisation, a penalty parameter (lambda) restricts the size of the estimated coefficients, thereby encouraging zero coefficients leading to the exclusion of redundant predictors. The non-constant model with highest lambda within one standard error of the model with minimum deviance was chosen to determine the best subset of predictors. As the algorithm is stochastic and may return different results at different runs, it was run 100 times and the result with the smallest deviance was selected.

4.2.8. COMPLEXITY TOPOGRAPHY

To investigate topographical areas of interest in discriminating responsiveness at moderate sedation, single-channel complexities (as computed before averaging to obtain LZSUM) were obtained. This is similar to K-complexity estimated by Sitt and colleagues (Sitt et al., 2014), but applied to the binarized time series. We employed a repeated measures ANOVA with level of sedation and responsiveness group as factors to probe for two-way interactions at individual channels. The resulting p-values were

controlled using the Benjamini-Hochberg procedure for false discovery rate (Benjamini and Hochberg, 1995). A significance threshold of 0.01 was used.

4.3. RESULTS

4.3.1. MEASURES OVERVIEW

Behavioural responsiveness and plasma drug concentration. Out of the 16 subjects, one became unresponsive during mild sedation and five more became unresponsive at moderate sedation. All subjects became again responsive during recovery. Plasma drug concentration had a high variance within levels: compared to the target levels of 0.6, 1.2 and 0 mg/L plasma drug concentrations, mean values were 0.47 (S.D.=0.2) mg/L during mild sedation, 0.93 (S.D.=0.25) mg/L during moderate sedation, and 0.3 (S.D.=0.09) mg/L during recovery (Figure 4.1A). Unresponsive periods had a mean plasma drug concentration of 0.99 (S.D.=0.35) mg/L, while responsive periods had a mean concentration of 0.5 (S.D.=0.26). There was considerable overlap in plasma drug concentration in the responsive and unresponsive groups (Figure 4.1A, B).

The number of responses and misses varied as expected with sedation level (Figure 4.1C, D). The mean reaction times (RTs), excluding misses, were 1.03s (S.D.=0.46) at baseline, 0.94 (S.D.=0.45) at mild sedation, 1.15 (S.D.=0.65) at moderate sedation and 0.87 (S.D.=0.36) at recovery. As expected, RTs were highest at moderate sedation (Figure 4.1E). Interestingly, baseline RTs were on average higher than RTs at mild sedation, and the overall lowest RTs occurred during recovery (Figure 4.1F), which likely reflected task habituation. This justifies the normalisation of RTs by the value at recovery later in the analysis. A Kolmogorov-Smirnov test (Massey, 1951) did not reject the hypothesis that the drug concentration and the RTs did not come from normal distributions ($p_{drug} = 0.6$, $p_{RT}=0.35$).

EEG measures. Two variants of temporal complexity were computed: LZSUM, which consists of the mean complexity of individual channels, and LZS, which operated on temporally concatenated values of single channels. The two versions of temporal complexity had a correlation of 0.99. Due to the very similar behaviour of these two measures, we henceforth only report temporal complexity as obtained from the LZS variant. The correlation between LZS and LZC was 0.6. A Kolmogorov-Smirnov test did not reject the null hypothesis that any of the complexity and integration measures were normally distributed, except for WSMI δ ($p_{LZT}=0.14$, $p_{LZC}=0.4$, $p_{WSMI\ \delta}=0.03$, $p_{WSMI\ \theta}=0.23$, $p_{WSMI\ \alpha}=0.15$, $p_{WSMI\ \beta}=0.54$).

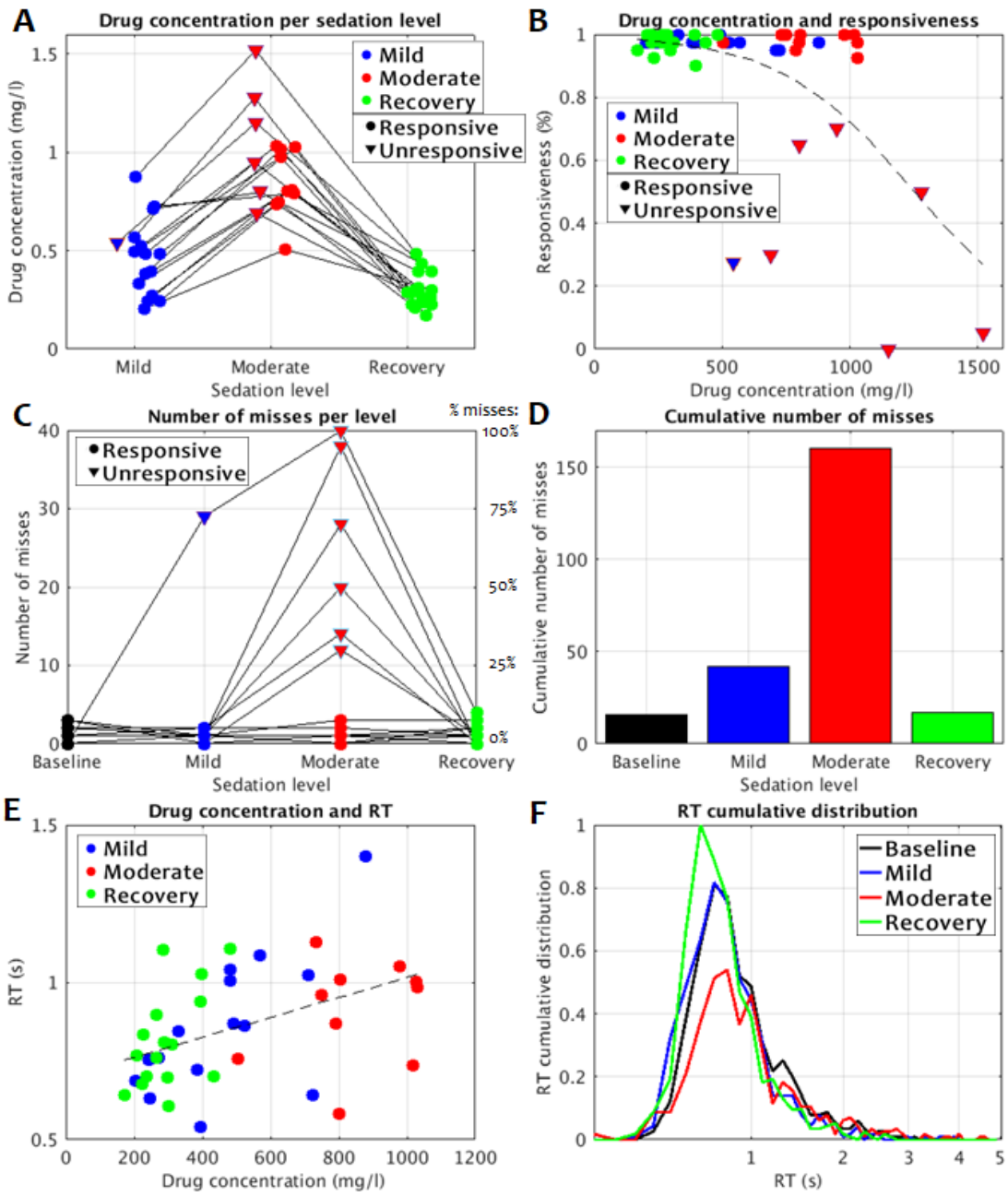

Figure 4.1 ***Behavioural and plasma drug concentration measures.*** *A: Individual plasma drug concentration at each sedation level. Values are jittered on the abscissa for visual clarity. B: Individual plasma drug concentration and responsiveness computed as the proportion of responses not missed during all sedation levels. C: Individual number of misses at each sedation levels. D: Cumulative number of misses per sedation levels. E: Individual plasma drug concentration and reaction time during all sedation levels. F: Cumulative distributions of reaction times at each of the four sedation levels.*

4.3.2. LEVEL ANALYSIS

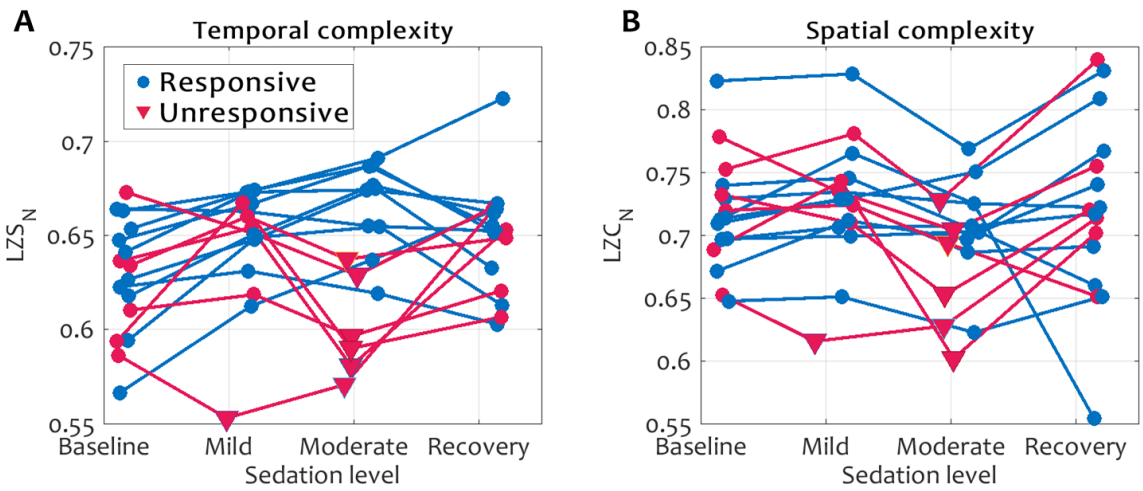

Figure 4.2 ***Temporal (A) and spatial (B) complexity of individual subjects at each sedation level.*** *Values are jittered on the abscissa for visual clarity. The subjects are classified as responsive or unresponsive based on the number of misses during the task at moderate sedation compared to the baseline period.*

Complexity. Although there was a high correlation between temporal and spatial complexity (Pearson's coefficient = 0.65), we found that the two measures tracked different aspects of sedation (Figure 4.2). A repeated measures ANOVA test with sedation level and responsiveness group as factors (see section 4.2.6 for details, including p-value correction) showed an interaction between responsiveness and sedation level for temporal complexity (F=5.82, p=0.002), but not in spatial complexity (F=1.65, p=0.2). Further, we used marginal means to explore group differences at moderate sedation, as well as the simple effect of state. For spatial complexity (Figure 4.2B), there was no difference between the responsive and unresponsive groups at moderate sedation (p=0.08), but there was a significant difference between baseline and moderate sedation across both groups (p=0.01). This suggests that spatial complexity is correlated with the increase in drug concentration, irrespective of responsiveness. Conversely, for temporal complexity (Figure 4.2A), there was a difference at moderate sedation between responsive and unresponsive groups (p=0.0001), driven by significantly increased temporal complexity in responsive (p=0.002) but not unresponsive (p=0.2) subjects. In contrast with spatial complexity, there was no significant difference between baseline and moderate sedation across both groups in temporal complexity values (p=0.7). This indicates that temporal complexity is able to track responsiveness, but not plasma drug concentration.

The above results were obtained after controlling for interference of spectral changes using phase randomisation as described in the Methods (section 4.2.4). In summary, the complexity values reported here are normalised first by the complexity of the shuffled initial string, and then by phase-randomised surrogate data. These measures hence correspond to the LZ_N measures in (Schartner et

al., 2017a). Importantly, we separately also confirmed that these effects were preserved when applying a notch filter to remove theta, alpha and beta oscillations. The effect sizes were smaller, but remained significant. Hence, we argue that this pattern of findings is driven by changes in signal complexity that cannot be fully explained away by changes in the oscillatory content of the EEG signal.

Information sharing. A repeated measures ANOVA test with sedation level and wSMI as factors showed no interaction for any of the four wSMI variants (Figure 4.3). However, there was a group difference between baseline and moderate sedation in wSMI δ (p=0.001) and wSMI θ (p=0.017). This was driven by a significantly lower information sharing in the unresponsive compared to the responsive group in both wSMI δ (p=0.023) and wSMI θ (p=0.02). This suggests that information sharing at these response frequencies potentially tracks responsiveness.

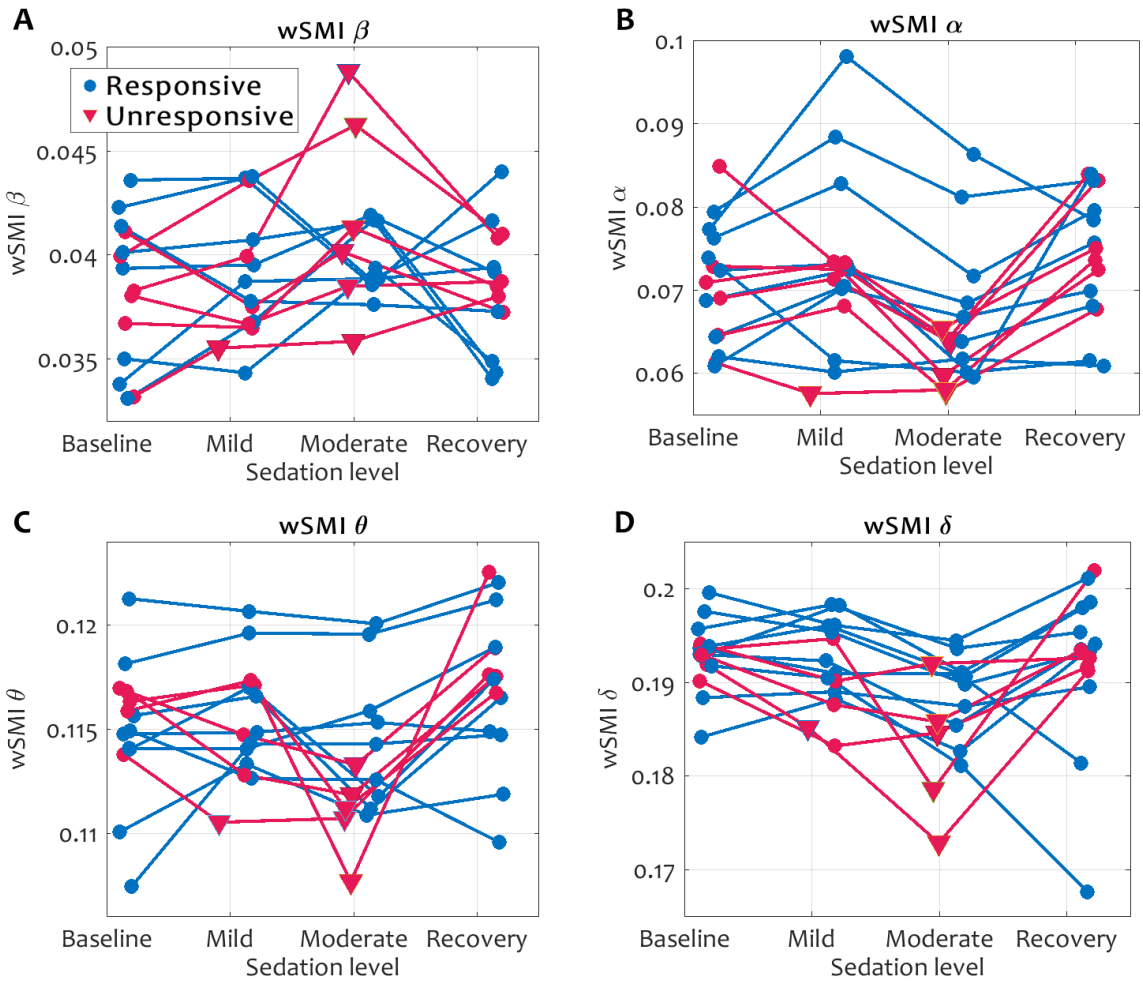

Figure 4.3 ***Trimmed-mean wSMI of individual subjects at four maximal response frequencies at each sedation level.*** *A: Beta. B: Alpha. C: Theta. D: Delta. Values are jittered on the abscissa for visual clarity. One outlier is excluded from the alpha and delta wSMI plots. Two outliers are excluded from the beta and theta plots. The subjects are classified as responsive or unresponsive based on the number of misses during the task at moderate sedation compared to the baseline period.*

4.3.3. OPTIMAL PREDICTORS

To assess the individual performance of each predictor, an individual GLM was first trained on each pair consisting of a predictor (complexity and wSMI values) and response (drug level, responsiveness and reaction time). Figure 4.4 shows the variance explained by each predictor, computed as the adjusted coefficient of determination. Drug concentration is best explained by spatial complexity and beta band information sharing. Responsiveness is best explained by temporal complexity and low-frequency (theta and delta) information sharing. No measures do particularly well in explaining reaction times.

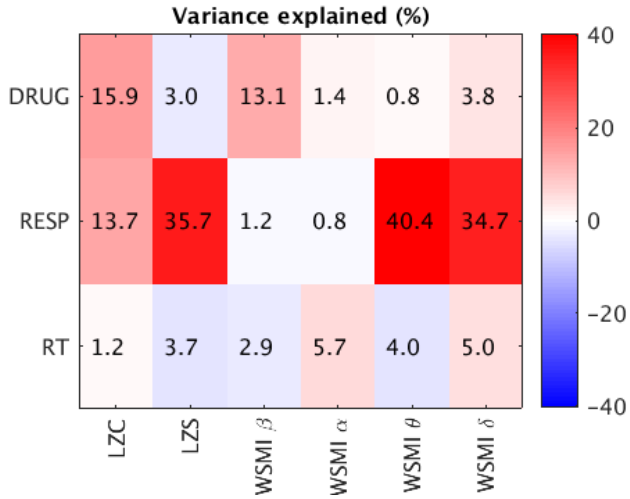

Figure 4.4 ***Variance of plasma drug concentration, responsiveness proportion and reaction time explained by complexity and information sharing predictors.*** *Variance explained is computed as the adjusted R^2 of the individually fitted regression model. Abbreviations are explained in Table 4.1.*

Abbreviation	Measure
LZC	Lempel-Ziv spatial complexity, obtained by concatenating data observation by observation.
LZS	Lempel-Ziv temporal complexity, obtained by concatenating data channel by channel.
LZSUM	Lempel-Ziv temporal complexity, obtained by averaging the complexity of individual channels.
WSMI β	WSMI with upper limit at beta frequency (tau=4 samples).
WSMI α	WSMI with upper limit at alpha frequency (tau=8 samples).
WSMI θ	WSMI with upper limit at theta frequency (tau=16 samples).
WSMI δ	WSMI with upper limit at delta frequency (tau=32 samples).
DRUG	Drug concentration level in blood.
RESP	Responsiveness, computed as the fraction of responses out of the total number of trials during a run.
RT	Reaction time, computed as the median per session in responsive sessions only.

Table 4.1 ***Abbreviations of variables used in the sedation analysis.***

To confirm the optimal subset of predictors for each outcome, we also performed Lasso GLM regularisation by predicting each of the three response variables using the full set of predictors. The following predictors, also shown in Figure 4.5, were selected as optimal:

- For drug concentration: spatial complexity (LZC), wSMI β and wSMI α;
- For responsiveness: temporal complexity (LZS), wSMI θ and wSMI δ;
- For RT: wSMI α and wSMI δ.

We therefore conclude that a combination of spatial complexity and higher-frequency information sharing best predict drug concentration, whereas a combination of temporal complexity and lower-frequency information sharing best predict responsiveness. RT is best predicted by alpha and delta information sharing.

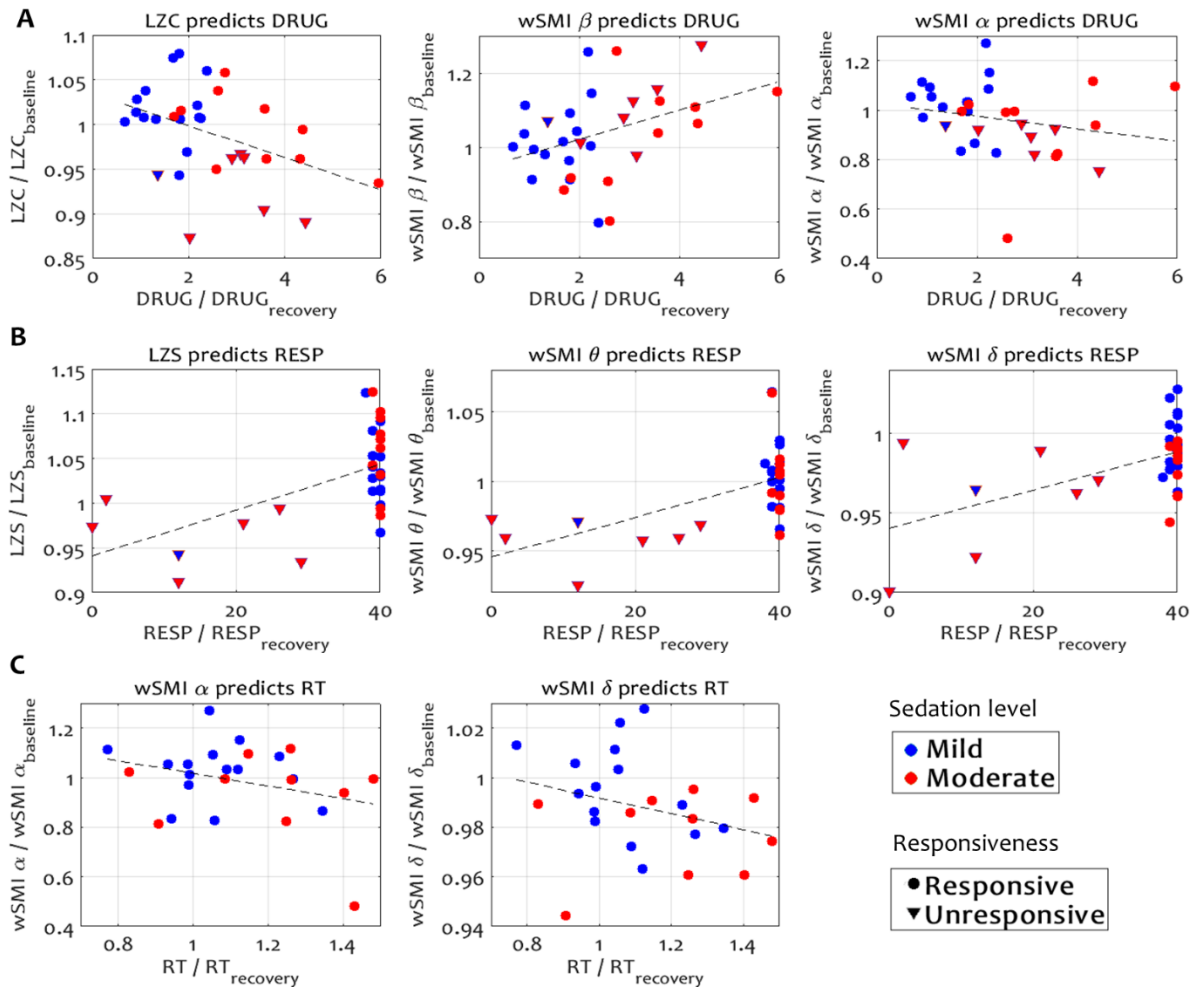

Figure 4.5 ***Best set of predictors selected using LASSO regularisation for plasma drug concentration (row A), responsiveness proportion (row B) and RT (row C).*** *Each subject contributes two points to each graph (from mild and moderate sedation). To increase point independence, individual values are normalised as described in the text. The dashed line shows the regression line fitted to each model (A, C: linear model; B: binomial model).*

4.3.4. COMPLEXITY TOPOGRAPHY

To understand what drives the difference in temporal complexity at moderate sedation in the case of responsive and unresponsive subjects, we investigated the topography of LZ complexity computed at single channels. In the baseline state, we found that complexity is highest in central areas and lower in posterior areas (Figure 4.6A). At moderate sedation, increased temporal complexity values are observed in the responsive group with pronounced anteriorisation (Figure 4.6B), whereas the

unresponsive group shows overall lower temporal complexity, except for a narrow central area (Figure 4.6C).

Figure 4.6D shows the difference between the change from baseline to moderate sedation in responsive and unresponsive participants. Highlighted channels show a significant interaction between level of sedation and responsiveness group after controlling the false discovery rate using Benjamini-Hochberg procedure and applying a significance threshold of 0.01. Single-channel complexity is significantly lower in frontal areas at moderate sedation in unresponsive subjects. Overall, this suggests that a gain in anterior complexity characterises preserved responsiveness during sedation, whereas the loss of anterior complexity is linked to unresponsiveness. This pattern was not observed in wSMI networks.

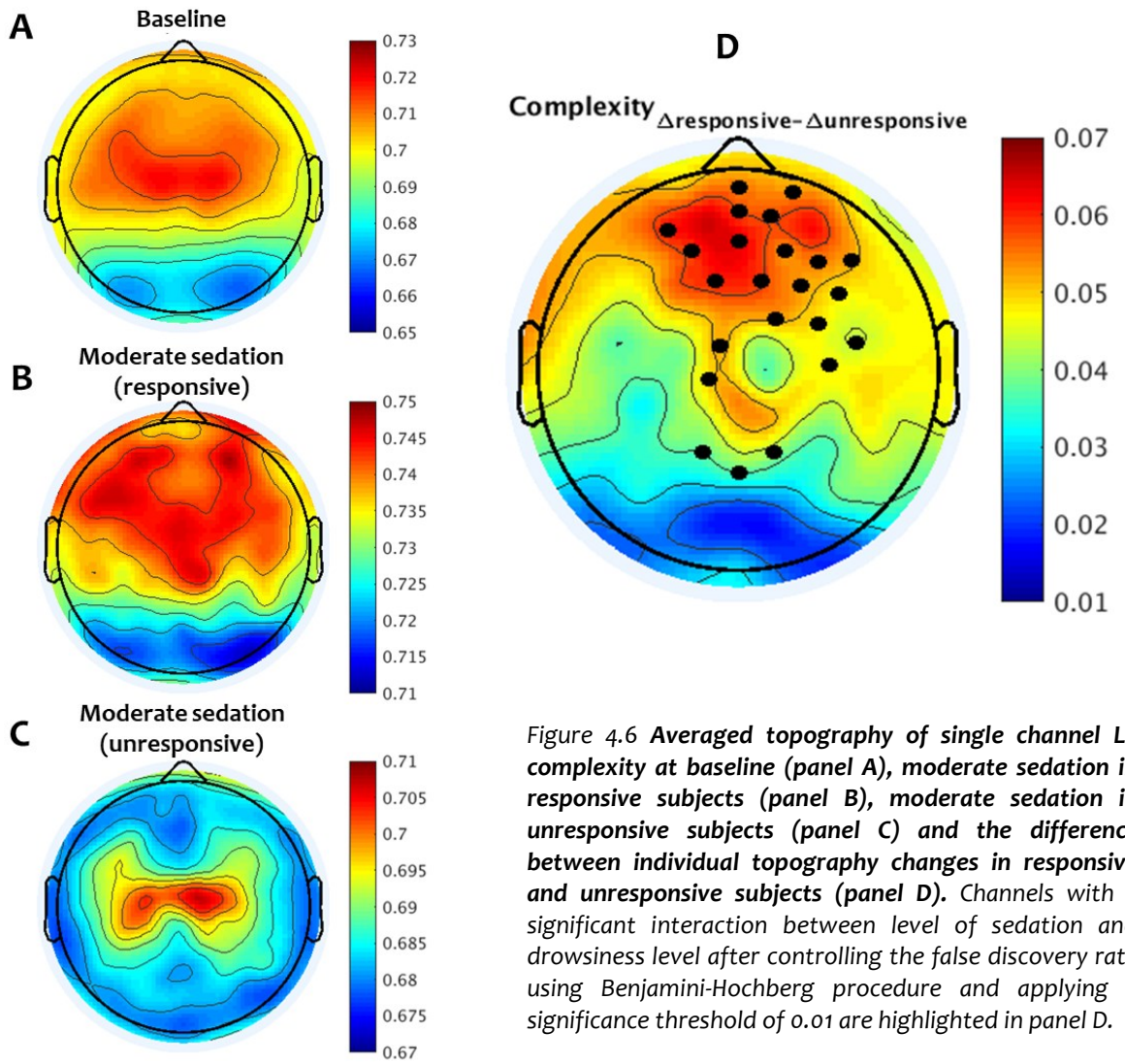

Figure 4.6 ***Averaged topography of single channel LZ complexity at baseline (panel A), moderate sedation in responsive subjects (panel B), moderate sedation in unresponsive subjects (panel C) and the difference between individual topography changes in responsive and unresponsive subjects (panel D).*** *Channels with a significant interaction between level of sedation and drowsiness level after controlling the false discovery rate using Benjamini-Hochberg procedure and applying a significance threshold of 0.01 are highlighted in panel D.*

4.4. DISCUSSION

4.4.1. SUMMARY

We investigated the change in information-theoretical measures of integration and differentiation during propofol sedation, motivated by the prediction of recent theories of consciousness that a decrease in information sharing and diversity of neural activity patterns leads to consciousness impairment. A perceptual discrimination task was used to assess responsiveness at baseline, mild sedation (target drug concentration 0.6 mg/L), moderate sedation (target drug concentration 1.2 mg/L) and recovery. Crucially, only 37.5% of the subjects lost responsiveness during moderate sedation, allowing us to discriminate responsiveness at similar drug levels. Plasma concentration of propofol and EEG measures of LZ complexity (temporal and spatial) and wSMI (at beta, alpha, theta and delta maximal response frequencies) were obtained at each sedation level. At light sedation doses, drug concentration is not sufficient to engender a loss of responsiveness. An analysis of the four sedation levels revealed that two types of LZ complexity tracked different aspects of sedation. Spatial complexity tracked sedation level indiscriminately of responsiveness, whereas temporal complexity differentiated between responsive and unresponsive subjects. We confirmed that these effects could not be fully explained by changes in spectral power. We also found that low-frequency information sharing discriminated between responsive subjects at moderate sedation. To further explore these findings, we applied regression analyses to find the best set of predictors for drug concentration, number of responses and median reaction time. We found that drug concentration was best predicted by spatial complexity and high frequency information sharing, whereas the number of responses was best predicted by temporal complexity and low-frequency information sharing. The best predictors for reaction times were alpha and delta wSMI, although they did not perform comparatively well. Finally, we found that complexity in anterior regions differentiated between responsive and unresponsive subjects during moderate sedation, with increased single-channel complexity in responsive subjects and suppressed single-channel complexity in unresponsive subjects.

4.4.2. UNRESPONSIVENESS AND CONSCIOUSNESS

It is important to clarify the relationship between the predictors employed in this study and the concept of consciousness. Here, we used two indices of alertness: reaction times, as a measure of efficiency in information processing, and the number of successful responses, as a measure of the capacity to fully process a stimulus. We acknowledge that the lack of responsiveness is not equivalent to a lack of consciousness (Boly et al., 2013a; Sanders et al., 2012). Propofol anaesthesia can involve conscious content in the form of unintentional awareness (Rampersad and Mulroy, 2005) or dreaming

(Brandner et al., 1997; Leslie et al., 2009). The dissociation of consciousness from responsiveness is a known issue for other states of impaired consciousness: for example, in disorders of consciousness, a minority of patients are able to produce brain activation similar to that of healthy individuals in response to task requirements despite the lack of an overt response, potentially suggesting the presence of covert conscious content (Cruse et al., 2011; Fernández-Espejo and Owen, 2013; Monti et al., 2010). Defining the level of consciousness based on a behavioural measure is therefore a limitation of the study. Nonetheless, the gradual loss of responsiveness, as well as the increased reaction times preceding unresponsiveness that we described here, indicate that responsiveness provides a good coverage of the early transition from consciousness to unconsciousness. In clinical practice, the absence of responsiveness is used routinely to assess the depth of anaesthesia as an acceptable proxy for consciousness (Callahan et al., 2017).

As expected, reaction times had a higher mean and variance during moderate sedation. On the other hand, reaction times (excluding misses) were faster at mild sedation than at baseline, and even faster at recovery. This can be explained as a combination of two factors: becoming more familiar with the task, as well as the effect of the drug at low doses being all-or-none, rather than gradual. It could be possible that the mild doses of sedative impaired the motor act of response itself, rather than the cognitive stage of the stimulus processing. The latter could have resulted in slower reaction times. The possibility of the motor response itself being impaired sooner than cognitive processing in early stages of losing consciousness is also supported by a study discussed in the previous chapter (Kouider et al., 2014), where it was shown that the lateralized readiness potential corresponding to the correct response can be detected in early stages of sleep despite the lack of responsiveness.

4.4.3. TEMPORAL AND SPATIAL COMPLEXITY AS DISTINCT NEURAL SIGNATURES OF RESPONSIVENESS AND PLASMA DRUG CONCENTRATION

Conceptually, how can there be a dissociation between the neural signature of responsiveness and that of plasma drug concentration? The individual response to estimated anaesthetic doses is variable across people (Araújo et al., 2017). Although plasma concentration level correlates well with changes in brain activity, not all of these changes are necessarily related to the ability to respond to external stimuli. By searching for the neural correlates of responsiveness, we focus specifically on a function normally required in full conscious states (Bor and Seth, 2012; Laureys, 2005; Laureys and Schiff, 2012). Indeed, a previous study from on fMRI data has confirmed that plasma drug level and responsiveness have a distinct set of optimal predictors obtained from BOLD functional connectivity networks (Barttfeld et al., 2015).

Here we report for the first time a dissociation between the spatial and temporal complexity of EEG data in predicting these two distinct aspects of sedation. Spatial complexity best tracks the plasma drug concentration, while temporal complexity performs better in predicting responsiveness. Importantly, it was ensured that these changes are not explained away by changes in power spectrum by using normalisation based on phase randomisation. It was also verified that the observed changes in these complexity measures remained even after applying notch filters on theta, alpha and beta bands.

By design, temporal and spatial complexity measure different aspects of the EEG signal: the former captures local diversity over time present in individual channels, while the latter captures global patterns across the whole topography, as well as the momentary relationship between channel values at individual timepoints. Our results suggest that the spatial diversity of scalp topography is reduced, i.e., it changes more slowly, as drug concentration increases. In contrast, the complexity of temporal dynamics in individual channels is reduced as responsiveness decreases. Interestingly, in the sedation level group analysis, we found that temporal complexity is in fact increased at moderate sedation in subjects who do not become unresponsive.

A similar dissociation between temporal and spatial complexity also exists in a previous study of propofol anaesthesia (Schartner et al., 2015). Here, Schartner and colleagues calculate LZ complexity using the same algorithm employed in this paper and on the same time window of 10 seconds. They administered a higher amount of propofol to their subjects and measured the level of anaesthesia using the Ramsay scale. The authors report obtaining similar results using the spatial and temporal versions of the LZ complexity, namely a decrease in this measure with the depth of sedation. However, a closer look at the values of temporal complexity they obtained (Figure S3 in their study) suggests that temporal complexity is in fact higher in mild sedation (propofol blood concentration 1.91±0.52 mg/L) than in wakefulness in 5 out of 7 subjects. This mirrors the increase in temporal complexity amongst the participants who remained responsive at moderate sedation in the current study. However, we report a smaller propofol blood concentration in the participants who remained responsive at this stage of our task (0.85±0.17 mg/L). In the remaining subjects, who became unresponsive, the concentration value was 1.06 mg/L. The study by Schartner et al. was performed on resting-state EEG data, therefore responsiveness cannot be compared. These results invite further investigation into the continuous complexity changes during mild sedation, in presence and in absence of a task. The same study (Schartner et al., 2015) also found that spatial complexity decreased linearly from wakefulness to mild sedation and into the loss of consciousness. In agreement with this, we also found that spatial complexity decreased with higher drug concentration.

In a different study on altered states of consciousness (Schartner et al., 2017a), it has been reported that temporal, but not spatial, complexity tracks the subjective experience during the administration of psychedelics. We did not collect subjective reports from out subjects, which is a limitation of our study. However, propofol has been previously reported to cause altered mental states in some patients (Balasubramaniam and Park, 2003; Brandner et al., 1997), suggesting that it might cause an altered state of consciousness at light doses where responsiveness is preserved. Temporal complexity is therefore a candidate neural signature for tracking neural processes related to altered states of consciousness, whereas spatial complexity is a better predictor for plasma drug concentration.

4.4.4. DECREASED FRONTAL COMPLEXITY IN UNRESPONSIVENESS

To understand what drives the change in temporal complexity at moderate sedation between responsive and unresponsive subjects, we examined the topography of single-channel LZ complexity. This is equivalent to the K-complexity used in other studies (Sitt et al., 2014). We described the baseline topography of complexity as anteriorised, with a peak in the central and anterior areas and lower values in the posterior area – an inverse map of typical spectral alpha power (Chennu et al., 2016a). At moderate sedation, the responsive group generally displays increased complexity, whilst preserving an anteriorised topography. In contrast, the unresponsive group shows a general drop in complexity and a loss of frontal complexity. The loss of frontal complexity in the unresponsive group could be related to the deactivation of the prefrontal cortex by propofol at moderate doses (Veselis et al., 2004). These results suggest that frontal complexity is essential in maintaining responsiveness during sedation.

4.4.5. LOW- AND HIGH-FREQUENCY INFORMATION SHARING AS NEURAL SIGNATURES OF RESPONSIVENESS AND DRUG LEVEL

To assess the dynamics of information integration during sedation, we calculated the wSMI with four distinct values for the temporal resolution parameter, which corresponded to maximal frequency responses at beta, alpha, theta and delta frequencies. The calculation of the wSMI includes a low-pass filtering at the calculated maximal frequency to prevent anti-aliasing artefacts. It has been shown that the wSMI generally peaks at a frequency close to the maximal frequency (King et al., 2013) (Fig. S2.h in their study), although there is a small degree of overlap with lower frequencies, especially as the temporal separation between considered samples becomes smaller. In this study, the results produced by the wSMI were overall noisy, so the findings should be interpreted with caution. Our results show that plasma drug concentration is best predicted by high frequency information sharing (alpha and theta), while responsiveness is best predicted by lower-frequency information sharing

(theta and delta). An interesting finding here is the increase in beta wSMI with deeper sedation, in contrast with delta, theta, alpha wSMI values, which show a simultaneous decrease. This emphasises the distinct functional roles of brain oscillations occurring at specific frequencies in the brain. This finding also mirrors the so-called beta-buzz phenomenon caused by propofol, where power at beta frequency increases at sedative doses (Gugino et al., 2001; Hashemi et al., 2017; Murphy et al., 2011; Purdon et al., 2013; Seifert et al., 1993). Whether there is a link between the increase in beta power and the increase in wSMI connectivity remains to be investigated further. Our results suggest that a higher drug concentration leads to increased information sharing at beta frequencies and decreased information sharing at frequencies below the beta band.

4.4.6. CONCLUSIONS

Our study supports the hypothesis that transitions of consciousness, such as sedation, can be tracked by the dynamics of neural activity differentiation and integration as measured using LZ complexity and wSMI. Intriguingly, we observed a dissociation between the neural correlates of two distinct aspects of sedation: drug level and responsiveness. For these two variables, we found two disjoint optimal sets of predictors: drug concentration was predicted by spatial complexity and high-frequency mutual information sharing, while responsiveness was predicted by temporal complexity and low-frequency information sharing. Most measure values decreased with deeper sedation, with two exceptions. Group-level temporal complexity increased in the responsive group at moderate sedation, and beta mutual information sharing increased with higher level of drug. Interestingly, the optimal subset of predictors for drug level and responsiveness included a simultaneous combination of both neural information integration and differentiation measures, indicating their complementary value in predicting the depth of sedation. Finally, we underlined the importance of frontal complexity in maintaining responsiveness during sedation.

Future work should seek to clarify several further questions related to these results. First, the relationship between temporal and spatial complexity should be better clarified conceptually, perhaps using simulated data to elucidate in detail the conditions in which a dissociation between the two might arise in a healthy or altered neural system. Secondly, it should be investigated whether spatial complexity, being a global measure, mirrors in practice any connectivity relationships between neural signals, thereby capturing, to some extent, the degree of integration occurring within the system. Secondly, the relevance of the frontal complexity module should be further explored, as part of ongoing debates about key areas of the brain whose functional alteration causes unresponsiveness, (Boly et al., 2017; Mashour and Avidan, 2017; Odegaard et al., 2017; Vijayan et al., 2013).

Overall, our study has both theoretical and clinical relevance. On the one hand, it provides evidence that supports and informs recent theories of consciousness, such as IIT. On the other hand, it is also of interest for clinical applications, with the prospect of enhancing patient monitoring during anaesthesia and sedation. The measures proposed here are easy to implement, computationally undemanding, and can be employed to track levels of sedation at a fine-grained scale during the loss of consciousness.

4.5. ACKNOWLEDGEMENTS

We thank Ram Adapa and Tristan Bekinschtein for performing the experiments, Jacobo Sitt for the wSMI code, and Andrés Canales-Johnson for feedback during early stages of this study.

CHAPTER 5

COMA

Having investigated the loss of consciousness in healthy adults due to natural and pharmacological causes – sleep and sedation –, we now move into the clinical realm. Here, theories of consciousness provide inspiration for the development of computational tools that can be used in the hospital to facilitate diagnosis and prognosis of patients with brain injury. This is particularly important considering the very high usual rate of misclassification in patients with disorders of consciousness (Schnakers et al., 2009). This chapter studies a cohort of acute comatose patients whose overnight EEG was recorded at the bedside in the intensive care unit during the first weeks after traumatic brain injury. Employing graph theory to measure the degree of integration and differentiation re-emerging in EEG networks early after injury, we find that increased variability in the characteristic path length of low-frequency networks predicts an eventual positive outcome. We also present the case study of a patient who, despite no behavioural improvement early after injury, showed remarkable early progress in the reconstruction of alpha connectivity and eventually fully recovered. This demonstrates how EEG tools inspired by theories of consciousness can be deployed at the bedside of patients with critical injuries to assist with clinical diagnosis and prognosis, and how such tools can inform on essential network configurations required for the re-emergence of consciousness.

Parts of this chapter have been presented as posters at the Association for the Scientific Study of Consciousness (ASSC) conference 2015 in Paris and at the Connectome Workbench 2015 in Cambridge.

5.1. INTRODUCTION

Consciousness is clinically defined as a state comprising both wakefulness and awareness (Baars, 1988; Laureys, 2005). Coma is a sign of brain injury so severe that no behavioural sign of consciousness is observable. It is usually caused by traumatic injury to the brain regions that sustain arousal in humans, notably the ascending reticular activating system (Edlow et al., 2013b; Laureys and Tononi, 2011). In coma, the brain does not sustain either wakefulness cycles or any observable awareness of the outside world (Laureys and Tononi, 2011). In contrast, in chronic disorders of consciousness (DoC), which include vegetative and minimally conscious states, patients show evidence of behavioural signs of

arousal and sleep-wake cycles. Vegetative patients are defined by a lack of awareness of the external world, whereas minimally conscious patients show transient signs of awareness. Patients who survive acute coma may either show progressive improvement in the degree of consciousness impairment up to full recovery, or reach a persistent vegetative or minimally conscious state (Laureys and Schiff, 2012).

Early prognosis of the potential outcome is of great importance to both clinicians and patient families in the decision-making process following injury, but this is particularly difficult in acute coma (Stevens and Sutter, 2013). Tools that are accurate and convenient to use at the bedside to diagnose and track the state of comatose patients would be highly valuable in intensive care units (Chennu et al., 2016b). The development of better behavioural (Arbour et al., 2016; Kalmar and Giacino, 2007; Schnakers et al., 2008; Stevens and Hannawi, 2016), biochemical (Herrmann, 2001; Vos et al., 2010, 2004) and neural (Bagnato et al., 2010; Chiappa and Hill, 1998; Edlow et al., 2013a) markers for coma prognostication is an active area of clinical research. Amongst these, the electroencephalogram (EEG) is a tool convenient to use in the bedside assessment of comatose patients. EEG methods proposed as predictors in coma prognosis include standard EEG visual inspection (Bagnato et al., 2010; Kaplan, 2004; Rumpl et al., 1983), spectral power (Bricolo et al., 1978), entropy (Gosseries et al., 2011) and evoked potentials, in particular the mismatch negativity (Amantini et al., 2005; Daltrozzo et al., 2007; Kane et al., 1996; Kotchoubey et al., 2001; Naccache et al., 2005).

Currently, behavioural measures that can be easily obtained at the bedside are commonly used for coma diagnosis. Standardised scales include the Glasgow Coma Score (GCS) (Jones, 1979) or the Coma Recovery Scale-Revised (CRS-R) (Giacino et al., 2004). However, behavioural scales are not always accurate. In DoC, misclassification rates as high as 43% have been reported when using standard clinical examination (Schnakers et al., 2009). By contrast, it has been established that neuroimaging can provide more insight into covert brain activity that can signal potential recovery. Vegetative patients who appear unaware of their surroundings may show patterns of brain activation surprisingly similar to healthy adults in response to verbal instructions, suggesting covert awareness of the outside world (Chatelle et al., 2012; Fernández-Espejo and Owen, 2013; Owen et al., 2006; Owen and Coleman, 2008). This suggests that brain activity could be a better and earlier predictor of recovery in coma compared to behavioural measurements.

Neuroscientific theories of consciousness (Oizumi et al., 2014; Tononi and Edelman, 1998) predict that a balance of functional integration and segregation across brain networks is essential for the brain in sustaining consciousness. EEG markers inspired by this prediction have been successfully used to characterise alterations of brain connectivity in coma (Achard et al., 2012), as well as in other DoC

(Chennu et al., 2017, 2014; King et al., 2013), anaesthesia (Chennu et al., 2016a) and sleep (Ferri et al., 2007; Larson-Prior et al., 2009). Brain connectivity patterns can be easily obtained from EEG recordings by applying measures of correlation or synchronicity (for example, consistent phase delays) on pairs of electrode signals (Sakkalis, 2011).

Graph theory can be used to quantify key properties of brain networks (Bullmore and Sporns, 2009; Sporns, 2010). Measures applied on neural connectivity networks can reveal the balance of segregation and integration across brain regions, as well as the overall small-worldness of the network, measured as the ratio between the two. It has been shown that the human brain has a small-world organisation (Bassett and Bullmore, 2006, 2016; Uehara et al., 2014). A study of graph-theoretical measures in a large cohort of DoC patients has shown that alpha connectivity and network centrality are useful predictors in clinical diagnostics (Chennu et al., 2017). Furthermore, graph theory applied to fMRI connectivity data of comatose patients has shown radically reorganized neural hubs of inter-modular information exchange compared to healthy adults, although other small-worldness properties were preserved (Achard et al., 2012). However, no study has yet characterised the early evolution of brain networks in acute comatose patients in relation to their eventual outcome. This could be a useful pursuit, given that acute comatose patients may display considerable changes in brain network reorganisation during the early post-injury period.

In this exploratory study, we set to investigate whether the temporal dynamics of connectivity networks constructed from high-density EEG data collected at bedside in acute stages of coma, following traumatic brain injury, can predict eventual clinical outcome in a small group of patients. In order to track the progression of the patient in this early stage, two recordings were obtained around one week and two weeks after injury, accompanied by the CRS-R score, followed by the CRS-R score outcome at two months after the traumatic event. The evolution of individual measures of brain connectivity from the first to the second recording was used as a predictor for the eventual outcome.

One challenge for the data collection in this study is the distorted architecture of the brain following the traumatic damage. This constrains the data collection as correct channel placement can be more difficult or even impossible, and will not necessarily correspond to the standard electrode locations on the scalp as described in healthy adults. Moreover, the spatial orientation of neural elements is important in the signal generation; for example, the apical dendrites of pyramidal neurons situated perpendicularly to the cortical surface are thought to contribute significantly to the EEG signal (Kirschstein and Köhling, 2009). Hence, brain geometry distortions may affect signal propagation itself. For this reason, in this study, we make no assumptions or analyses concerning specific brain locations that have been associated with various cognitive functions in healthy adults, but rather investigate

the configuration of function networks in a geometric sense. We mention, however, long- in contrast with short-range connectivity, as we expect the former to be informative of the ability of neural populations to communicate across brain areas in spite of the architectural changes caused by trauma, and thus to be a positive marker with respect to the prognosis.

We acknowledge that traumatic brain injury leads to a highly disrupted neural configuration that is particular to each individual case. For the scope of this small-sample study, we did not focus on the clinical aspects of the individual injuries or on the structural brain damage, as there was not sufficient statistical power available to incorporate these factors into our analysis. Instead, by tracking in detail the dynamical change in brain network measures over time within each patient, we obtain individual metrics that can predict eventual improvement. This generalised approach allows us to focus on creating a general computational pipeline that could be easily deployed at the bedside for real-time monitoring of a patient, an important goal in the clinical care of DoC patients (Chennu et al., 2016b). Moreover, this pipeline aims to deal efficiently with the relatively high level of artefacts inevitable in the patient EEG recordings, so we aimed for a limited computational cost and no manual intervention, with conservative thresholds for data quality control. In this sense, another contribution of this work is methodologically motivated.

5.2. METHODS

5.2.1. SUBJECTS AND DATA COLLECTION

Of 17 patients included in the study, two overnight EEG recordings obtained at the bedside were available for 11 non-sedated patients (5 female; mean age 37.8, S.D. = 15.5) in acute coma after traumatic brain injury. Table 5.1 shows a description of individual patient information (Numbering discontinuities are due to other patients initially recruited in the study, but where a second recording session was not possible due to clinical reasons.). The first recording session took place for all patients after a mean of 8.5 days (S.D. = 1.8) following the injury and a second recording was obtained after a mean of 15.6 days (S.D. = 2.3) following the injury. At the time of both recordings, the patients had not been sedated for at least 48 hours. CRS-R scores, as measured by a clinician, were stored for each recording session, as well as the outcome at two months after injury. The CRS-R has been validated as a reliable measure across tests performed in post-comatose patients (Bodien et al., 2016; Schnakers et al., 2008).

Where possible, we aimed to obtain continuous recordings during the whole night starting around 18:00 hours. The mean length of the raw recordings was 10.9 hours (SD 4.4 h; see also Figure 5.3). Recording length varied across patients for clinical reasons, but no recording was shorter than 1 hour.

The EEG data was acquired continuously using EGI's Geodesic EEG Net Amps 300 system (Electrical Geodesic Inc., Oregon, USA) with a gel-based sensor net of 128 electrodes, referenced to the vertex, at a sampling rate of 250 or 500 Hz.

The recordings were performed at the Addenbrooke's Hospital in Cambridge, UK. Ethical approval was provided by the National Research Ethics Service (National Health Service, UK; REC number: 974/290). All patients were recruited and managed according to Addenbrooke's Neurosciences Critical Care Unit cerebral perfusion management algorithm (Menon, 1999) and informed consent was acquired from the families of all patients and their medical teams before recording.

Patient number	Age	Gender	Final CRS-R	Lesion site
2	55	M	20	Right frontoparietal hematoma + intraventricular haemorrhage
3	37	F	5	Diffuse white matter intensities
5	19	F	17	Left temporofrontal extradural hematoma + left hemispheric subdural hematoma
6	31	F	7	Bilateral frontoparietal SDH + Diffuse white matter intensities
7	53	F	6	Right frontoparietal subdural hematoma + Left temporal extradural hematoma
10	21	F	7	Bilateral frontoparietal subdural hematoma
11	24	M	11	Bilateral frontoparietal hematoma + intraventricular haemorrhage
13	54	M	5	Diffuse white matter intensities
15	29	M	9	Left temporofrontal extradural hematoma + left hemispheric subdural hematoma
16	42	M	16	Diffuse white matter intensities
17	28	M	6	Right convexity subdural hematoma + bilateral haemorrhagic cortical contusions

Table 5.1. ***Age, gender, outcome and lesion details of individual comatose patients.***

5.2.2. DATA PRE-PROCESSING

In addition to an investigational analysis of the dataset, a central aim of this research was to develop an analysis pipeline that could eventually be deployed to tracking the state of the patient in real-time. Pre-processing scripts were facilitated by the EEGLAB toolbox (Delorme and Makeig, 2004) for MATLAB. First, EEG channels located on the neck, cheeks and forehead were discarded in order to minimize muscular artefacts, leaving 92 channels for further processing. All data was resampled at

250 Hz if recorded with higher sampling, and filtered between 0.5 and 25 Hz using a finite impulse response filter with a Hamming window (Blackman and Tukey, 1958). The data was then divided sequentially into 10-second epochs and channel means were subtracted from each epoch for baseline correction. For data cleaning, the standard deviation of every channel at every epoch was computed. A channel was considered too noisy if its standard deviation exceeded 250 microvolts, a threshold established by visual inspection with the aim to remove artefacts while keeping as much data as possible. With the same reasoning, an epoch was rejected if more than 10% (9) of the channels were too noisy; otherwise, noisy channels were interpolated. Finally, the data was re-referenced to the common average of all electrodes at each time point, resulting in 91 channels in each dataset. With a large enough number of electrodes (in practice, more than 64), the common average reference closely approximates the underlying signal (Nunez and Srinivasan, 2006).

Figure 5.1 illustrates the processing pipeline for one subject, as described in the following sections.

5.2.1. POWER SPECTRA

Power spectra were computed for each epoch, at frequencies between 1 and 13 Hz in steps of 0.1 Hz, using Hann windows (Blackman and Tukey, 1958). Computations were performed using the Fieldtrip toolbox (Oostenveld et al., 2011). Power spectra analyses are reported at channel Pz (E62), which has been previously used in other EEG studies on comatose patients (e.g. Kane et al., 1996) and is a suitable location for detecting frequency rhythms originating in several regions across the scalp, including occipital alpha. It was checked that the results remained qualitatively the same for an average over all electrodes, as well as for predefined regions of interest when analysing occipital channels for alpha power, central channels for theta power, and frontal channels for delta power, as used in similar studies (Chennu et al., 2014). For every epoch, the total power at channels of interest in the delta (1-4 Hz), theta (4-8 Hz) and alpha (8-13 Hz) frequency bands was summed and their relative contribution to the total power between 1 and 13 Hz was stored. Previous literature has shown these to be the most informative frequencies in disorders of consciousness, as the EEG is slowed down before full recovery (Chennu et al., 2014; Schiff et al., 2014; Sitt et al., 2014). To summarise the results, the overall median and standard deviation of the relative contributions of the three bands over time was used in subsequent analyses, as described below.

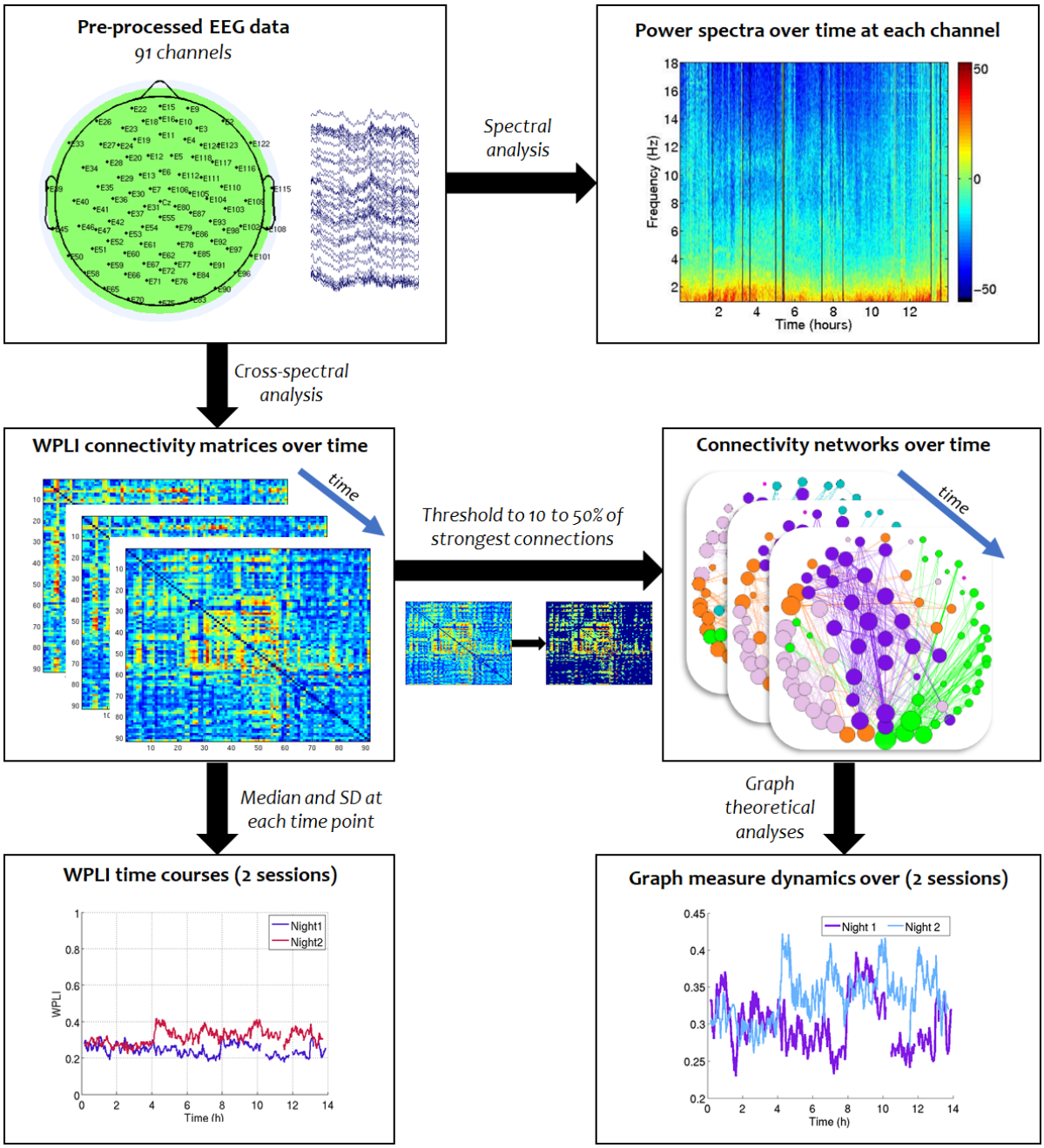

Figure 5.1. ***Illustration of the analysis pipeline.***

5.2.2. CONNECTIVITY NETWORKS

To build connectivity networks, we employed the weighted phase lag index (WPLI) (Vinck et al., 2011), a measure based on the observation that long-distance synchronization of oscillatory activity in the human brain, in particular its phase coupling, is fundamental to information processing (Sauseng and Klimesch, 2008). However, a number of connectivity measures are prone to overestimating connectivity due to volume conduction, which results in spurious correlations between signals

affected by the same source (Nunez and Srinivasan, 2006). The WPLI is based on the phase lag index (PLI) (Stam et al., 2007), which is designed to ameliorate the problem of volume conduction by ignoring zero phase differences, as these are likely caused by volume conduction. The PLI is computed as an average over the signs of nonzero phase differences between two signals. The weighted phase lag index (WPLI), has been introduced as an improvement to the PLI (Vinck et al., 2011). In the calculation of the WPLI, the signs of the phase differences are weighted by their absolute magnitude, such that they are ignored at 0 and 180° and maximally weighted at 90 and 270°. It has been verified that the WPLI, though conservative in its estimation, avoids misidentifying volume conduction as true connectivity (Cohen, 2014) and, furthermore, that it is a robust measure for network analyses when applied at different recording times in the same subject (Bassett et al., 2011).

The calculations were performed using a debiased estimator of the squared WPLI that corrects for bias in small sample sizes, as implemented by the Fieldtrip toolbox. Debiased WPLI was estimated at frequencies between 1 and 13 Hz in steps of 0.1 Hz, for each pair of channels. We employed a sliding-window analysis, where WPLI connectivity matrices of size 91 x 91 were computed on windows of 60 10-second epochs, with an overlap of 10 epochs between windows. As in the previous cleaning step, a conservative threshold of 10% (6) epochs was set for all subjects. The WPLI was only computed over the non-rejected epochs of a window if the number of rejected epochs was less than or equal to the threshold. Finally, to select the highest observed connectivity value in each time 10-minute window, the maximum WPLI value was selected at each pair of channels within each of the three predefined frequency bands: alpha (8-13 Hz), theta (4-8 Hz) and delta (1-4 Hz).

5.2.3. CONNECTIVITY TIMECOURSES

The above calculation generated a timecourse of connectivity matrices over each of the two recording sessions for each patient. As a first-order approach to assess the overall strength and variability of overnight connectivity in individual recordings, four measures were computed for each recording, at each frequency band of interest. Median connectivity over all channel pairs was computed at each 10-minute window. Then, to summarise the overnight timecourses, the overall median and the standard deviation of these medians was computed. Further, to assess variability in connectivity levels, the standard deviation over WPLI values was computed in each window, and the overall median of these standard deviation values was stored. Finally, to assess the variability within individual connections, the standard deviation of every channel pair was computed over all windows and the median of these standard deviations was recorded. Similar timecourses were also obtained for the graph-theoretical measures, as described below.

5.2.4. GRAPH THEORETICAL MEASURES

A matrix of WPLI values obtained as described above can be seen as the adjacency matrix of a weighted undirected graph spanning over the scalp, composed of electrodes located in sensor space as nodes and connectivity values as weighted edges. As a higher-order approach to assess network characteristics, graph-theoretical measures were computed on WPLI matrices at each frequency band for a range of thresholds keeping between 10 and 50% (step size 2.5%) of the strongest connections. For each measure described below, its median across all thresholds was used in further calculations. This approach has commonly been used in other brain connectivity studies at various scales (Achard and Bullmore, 2007; Chennu et al., 2014; Lynall et al., 2010), as it provides a balance between a too high connection density, which would result in random network characteristics by including weak (possibly spurious) connections, and a too low density, which might lead to a disconnected network. Above the threshold, graph weights were unchanged, in order to better preserve the original architecture of the network, avoiding potential false shortcuts that could be created in the graphs by binarising them (Monti et al., 2013; Rubinov et al., 2009).

Computations were performed using the Brain Connectivity Toolbox by Rubinov and Sporns (2010) (also see this reference for more computational details on each measure). The measures are chosen to reflect segregation, integration and centrality properties of the networks.

To assess the degree of segregation into functional modules in each network, the clustering coefficient of each node and the network modularity were calculated at each time window. The clustering coefficient (Watts and Strogatz, 1998), measures the number of triangles formed around a node through connections between its neighbours, thus describing local connectivity. In its adaptation for weighted graphs, the geometric mean of each triangle is used to compute an average clustering coefficient (Onnela et al., 2005). The median clustering coefficient and its standard deviation over all nodes was computed for each time window.

Modularity (Newman and Girvan, 2004) was also computed as a mesoscale property showing how well the network can be partitioned into non-overlapping communities (modules). The Louvain heuristic (Blondel et al., 2008) uses a greedy optimization technique to minimize the modularity of the network by gradually selecting nodes to include into modules. Modularity has been shown to be significantly elevated in the human connectivity map compared to surrogates with realistic topological constraints (Samu et al., 2014).

Macroscale integration across the network was assessed using the characteristic path length (Watts and Strogatz, 1998), which is the average shortest distance between all pairs of nodes in the network.

The distance between two nodes is given by the minimum sum of weights describing a sequence of connected nodes starting and ending at the two given nodes. As higher values in the WPLI matrix represent shorter distances, the inverse of the weighted connectivity matrix was used for to compute the path length.

Centrality (Freeman, 1978) of the nodes in the networks was measured using the betweenness and the participation coefficient of all nodes at each time step, thus quantifying the evolution of nodes acting as hubs that facilitate information exchange across modules in the network. The betweenness of a node (Brandes, 2001) represents the proportion of shortest paths in the network which contain it, whereas the participation coefficient (Guimerà and Amaral, 2005) measures the diversity of connections from a node to different modules. The median and standard deviation of these measures were computed across all nodes at each time window.

As graph measures are unaware of the spatial embedding of the network, the modular span (Chennu et al., 2014) was also computed for the purpose of measuring the median topographical distance spanned by a module relative to its number of nodes. Modular span was introduced as a successful method of discriminating EEG network topographies of DoC patients compared to healthy adults.

Graph analyses were concluded by computing the small-worldness of the network, calculated as the ratio between the median clustering coefficient over all nodes in a network and the characteristic path length of the network. The original measure of small-worldness of a network (Humphries and Gurney, 2008) is computed by normalizing the above ratio by the same ratio computed in a matching random network (a network with the same number of nodes and edges, but randomly distributed). As, in this study, we compared the evolution of network properties from the first to the second session by taking a ratio (as described below), the denominator produced by the random network would cancel out in the final computation, so the small-worldness measure did not require normalisation.

5.2.5. OUTCOME CORRELATIONS

To correlate these measures with the patient outcomes, every measure was first summarised as a single value representing a median or a standard deviation during one recording session. The median was used instead of the mean in order to minimize the effect of outliers, in particular for sessions with a smaller number of time windows. Therefore, for each patient, two values were available for each measure, corresponding to two recording sessions in the acute stage of the injury, taken around a week apart. These two values were used to compute the proportional change of a measure from the first to the second recording session, as an indicator of early progress. The proportion change was used as a predictor for the outcome at two months, as assessed by the CRS-R score. This approach

effectively tested whether changes in EEG parameters over two recording sessions a week apart shortly after injury could predict longer term outcome.

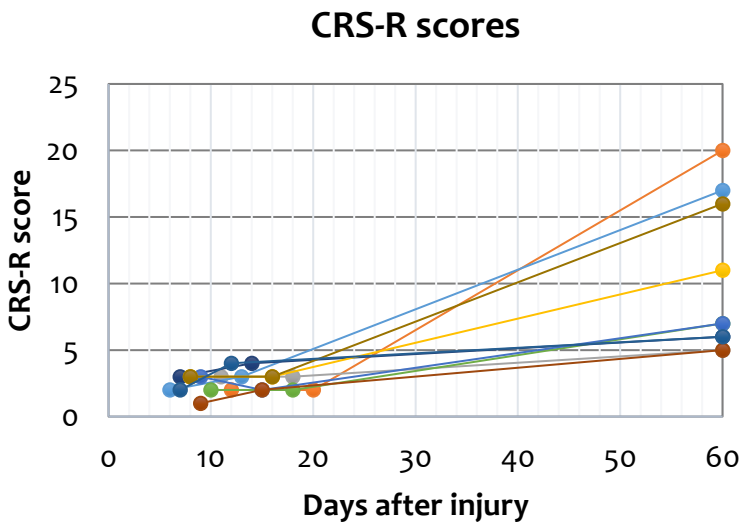

Figure 5.2. ***CRS-R scores of the comatose patients at the three time points measured post-injury.***

Assuming a linear relationship between the proportion change in a measure and the patient outcome, the p-value and R^2 coefficient of determination were calculated for the linear regression fit of the proportion change for each measure as predictors of the CRS-R score at two months of each patient. Due to the small sample size, no outliers were removed before performing the correlations. The statistics were corrected for multiple comparisons using the Bonferroni-Holm method (Holm, 1979) at an alpha level of 0.05, to correct for all tests performed. For added robustness, the results were also checked using Spearman's rank correlation.

5.3. RESULTS

5.3.1. OVERVIEW

The behavioural analysis showed that there was no correlation between the change in the CRS-R score from the first to the second session and the CRS-R outcome after two months, as shown in Figure 5.2. At the time of the EEG recordings, which were obtained one to three weeks after injury, all CRS-R scores were below 5, while the scores at two months were distributed from 5 to 20.

The EEG data displayed a high level of noise in many of the subjects. Figure 5.3 depicts the amount of EEG data available for each patient, detailing the number of channels interpolated for each subject and the amount of data rejected.

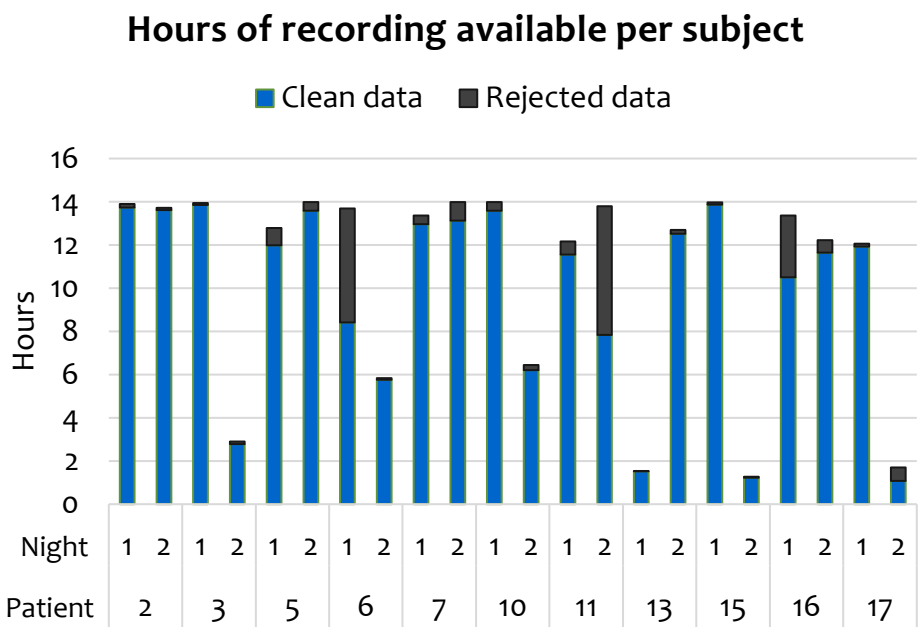

Figure 5.3. ***Data availability for individual subjects and nights.*** *Dark grey denotes data rejected due to more than 10% of the channels being classified as noisy. The data cleaning procedure is described in section 5.2.2.*

An overview of the correlation between individual measures and behavioural outcome is given in Table 5.2 for all measures that were tested. The measures that performed best are highlighted. In this patient cohort, it was found that the change in variability in characteristic path length in delta networks in the acute phase of the injury was a significant predictor of eventual behavioural outcome, which remained significant after correction for multiple comparisons ($p = 0.004$). However, the sample size was too small to allow sufficient statistical power for generalisable results, so these findings should be interpreted with caution. In the next sections, an exploratory account of individual measures that performed best is presented.

Notably, in this cohort, the patient with the best outcome (CRS-R = 20) was often found to be an outlier among the group of comatose patients. This patient (referred to below as patient P2) displayed unique brain network changes from one to two weeks after injury, despite no behavioural signs of improvement. As this patient was the most interesting in this cohort, P2 is depicted below as a case study for measure dynamics indicating an evolution towards recovery. To demonstrate the dynamics of connectivity networks in this patient, a video comparing the modular decompositions of theta networks during the first and second recording session is available online at https://vimeo.com/124935436.

Measure	Linear model p-value	Linear model R^2	Spearman's p-value	Spearman's Rho
SD Characteristic path length - delta	**0.00006**	**0.84594**	**0.00568**	**0.7689**
Median betweenness - alpha	**0.02072**	**0.4655**	**0.04687**	**0.60871**
SD betweenness - delta	**0.03556**	**0.40404**	**0.02889**	**0.65448**
SD betweenness - alpha	**0.04357**	**0.37952**	**0.0233**	**-0.67278**
SD Median clustering - delta	**0.07273**	**0.31443**	**0.00129**	**0.83755**
Median participation - delta	0.0844	0.29473	0.09973	0.52175
Modular span - theta	0.09444	0.27962	0.12627	0.48971
Modularity - delta	0.09504	0.27877	0.2194	-0.40276
SD Small-world-ness - delta	**0.10411**	**0.26638**	**0.01039**	**0.73228**
SD Modular span - theta	0.10416	0.26631	0.1346	0.48056
Small-world-ness - theta	0.10491	0.26533	0.14773	0.46683
Median participation - alpha	0.10953	0.25943	0.38987	-0.28834
SD clustering - theta	0.12255	0.24396	0.1346	0.48056
Characteristic path length - theta	0.12365	0.24272	0.11061	-0.50802
SD participation - theta	0.13428	0.23127	0.22518	0.39818
Modularity - theta	0.14022	0.22523	0.24305	-0.38445
Median clustering - theta	0.14239	0.22309	0.1346	0.48056
SD Modular span - alpha	0.14317	0.22232	0.24305	0.38445
Modular span - alpha	0.18507	0.18623	0.44681	0.2563
SD participation - alpha	0.19279	0.18046	0.32954	0.32495
Median betweenness - delta	0.22251	0.16023	0.16646	-0.44852
SD Median betweenness - alpha	0.25372	0.14177	0.5163	0.21968
SD Modular span - delta	0.29102	0.12266	0.43845	0.26088
Median betweenness - theta	0.38807	0.08374	0.35155	-0.31122
Characteristic path length - alpha	0.40294	0.07884	0.35905	-0.30664
SD Modularity - delta	0.44581	0.06597	0.91489	0.03661
SD Median participation - alpha	0.44965	0.0649	0.68726	0.1373
Median participation - theta	0.45151	0.06439	0.07138	-0.56294
Small-world-ness - alpha	0.56658	0.03783	0.6283	0.16476
SD Median clustering - alpha	0.56977	0.03723	0.65756	-0.15103
Modularity - alpha	0.58514	0.03439	0.97869	-0.00915
SD Median betweenness - theta	0.58939	0.03363	0.86205	0.0595
SD betweenness - theta	0.60116	0.03158	0.481	0.23799
SD Characteristic path length - theta	0.62904	0.02703	0.66742	-0.14646
SD Median participation - theta	0.63394	0.02628	0.9255	0.03204
Median clustering - alpha	0.63625	0.02592	0.79932	0.08696
SD clustering - alpha	0.64805	0.02417	0.84104	0.06865
SD Small-world-ness - theta	0.67131	0.02092	0.70729	0.12815
SD clustering - delta	0.68392	0.01928	0.90429	-0.04119
SD Small-world-ness - alpha	0.68858	0.01869	0.53436	-0.21053
SD Median betweenness - delta	0.69426	0.01798	0.64776	-0.15561
Characteristic path length - delta	0.74784	0.01206	0.88313	-0.05034
SD Characteristic path length - alpha	0.76306	0.01062	0.35905	-0.30664
Median clustering - delta	0.76469	0.01047	0.94675	-0.02288
Small-world-ness - delta	0.76846	0.01013	0.93612	-0.02746
SD Modularity - theta	0.79004	0.00829	0.28781	0.35241
SD Modularity - alpha	0.83026	0.00538	0.91489	-0.03661
Modular span - delta	0.89293	0.00213	0.53436	-0.21053
SD Median participation - delta	0.89742	0.00195	0.80971	-0.08238
SD participation - delta	0.96036	0.00029	0.481	-0.23799
SD Median clustering - theta	0.9894	0.00002	0.85153	-0.06407

Table 5.2. ***Linear regression and Spearman correlation p-values and coefficients (uncorrected).*** *Rows are ordered by p-value in the linear model. 'SD' refers to the standard deviation of a measure. Measures where a p-value is below 0.05 are highlighted.*

5.3.2. ALPHA NETWORKS

Alpha networks are fundamental in healthy brain functioning (Chennu et al., 2014; Klimesch, 2012). We found a uniquely strong increase in alpha band power from the first to the second night in P2, the only patient who progressed to a full recovery after two months (Figure 5.4). Graph theory analysis also revealed an interesting evolution of nodes acting as central connectivity hubs in P2's alpha network: the standard deviation of betweenness at each time window is much higher during the first session, indicating a high functional diversity of individual nodes; it evolves to a higher median and a smaller standard deviation in the second night, indicating that more nodes have the same degree of participation in shortest paths in the graph (Figure 5.5). Here the decision to use the median as a summary measure is crucial, as the nodes with higher variability would disproportionately influence the mean during the first session. Overall, P2 is an outstanding subject considering the dynamics of his alpha networks before any behavioural sign of improvement was shown, suggesting that early alpha network reconstruction is highly beneficial for recovery.

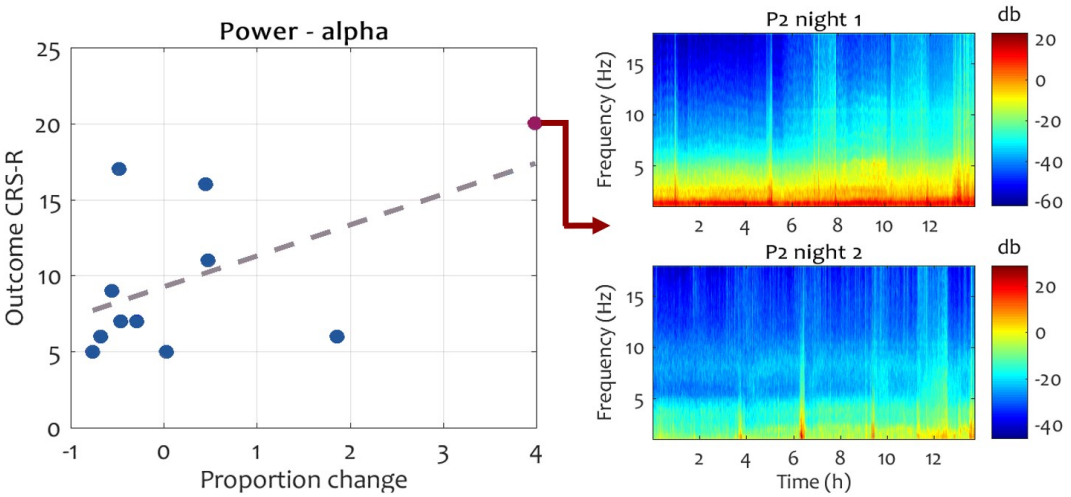

Figure 5.4. ***Alpha power changes.*** *(Left) Proportion change in alpha contribution to total power at channel Pz (E62) from the first to the second session in all patients. The dashed line represents the linear fit. (Right) Power spectra over the whole recording sessions (top: first session; bottom: second session) for P2, averaged across all channels. Black vertical lines are due to rejected data epochs.*

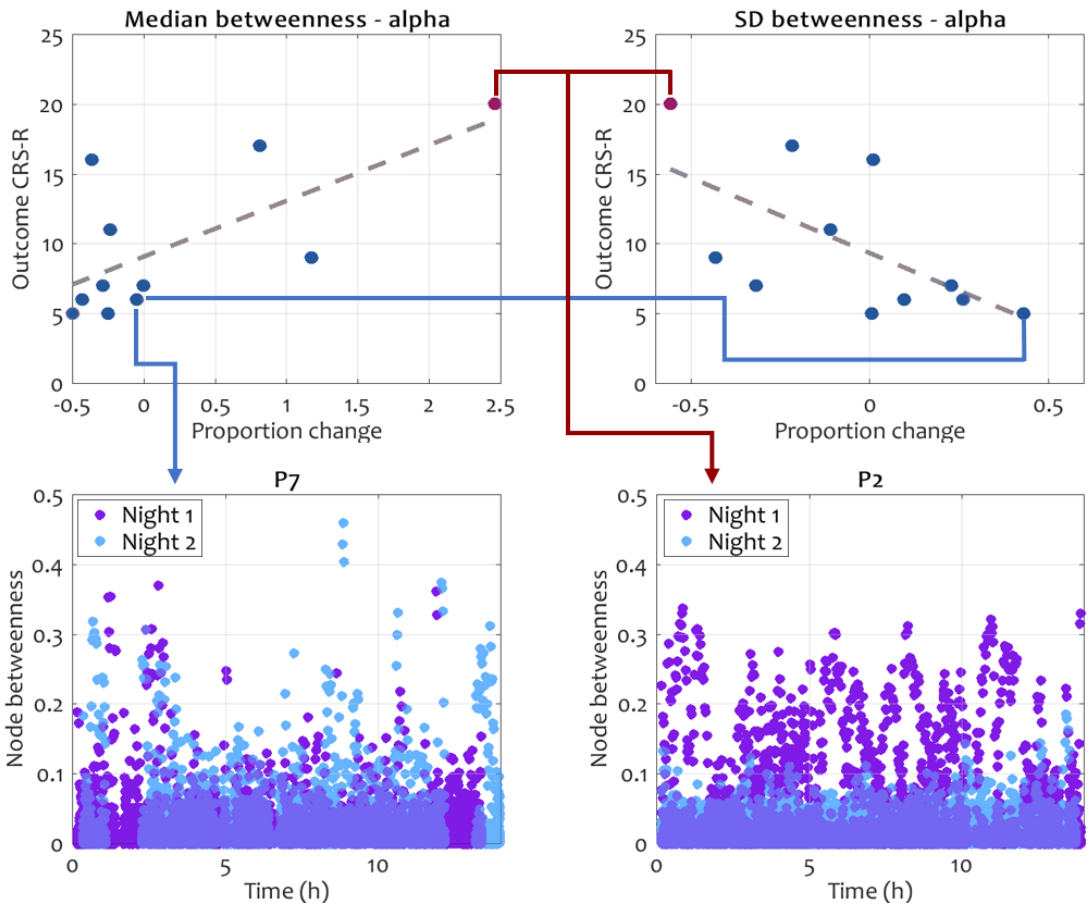

Figure 5.5. ***Node betweenness of alpha networks.*** *(Top row) Proportion change from first to second session in median node betweenness (left) and the variability of median node betweenness (right) across time in alpha networks in all patients. Dashed lines represent linear fits. (Bottom row) Overnight dynamics of node betweenness in alpha networks during the two sessions in two selected patients. All nodes are shown. P2 shows an interesting evolution: although the extreme betweenness values from the first session are attenuated, the median during the second session is in fact higher and there is less variability in the second session (right). On the other hand, more variability is observed in a patient with a poor outcome (left). Discontinuities are due to rejected data epochs.*

5.3.3. THETA NETWORKS

In patients with poor outcome, there was a decrease from the first to the second recording in median connection strength in the theta band. This was also the case for the median standard deviation of individual connections (Figure 5.6). In patient P2, frontoparietal theta connections had the highest variability. Furthermore, graph theory measures in theta band show that, in this patient, functional modules of synchronised activity comprise longer range neural modules in the second session compared to the first (Figure 5.7). Overall, this is also reflected in the median standard deviation of clustering in theta networks, with a poor outcome observed in patients with lower standard deviation.

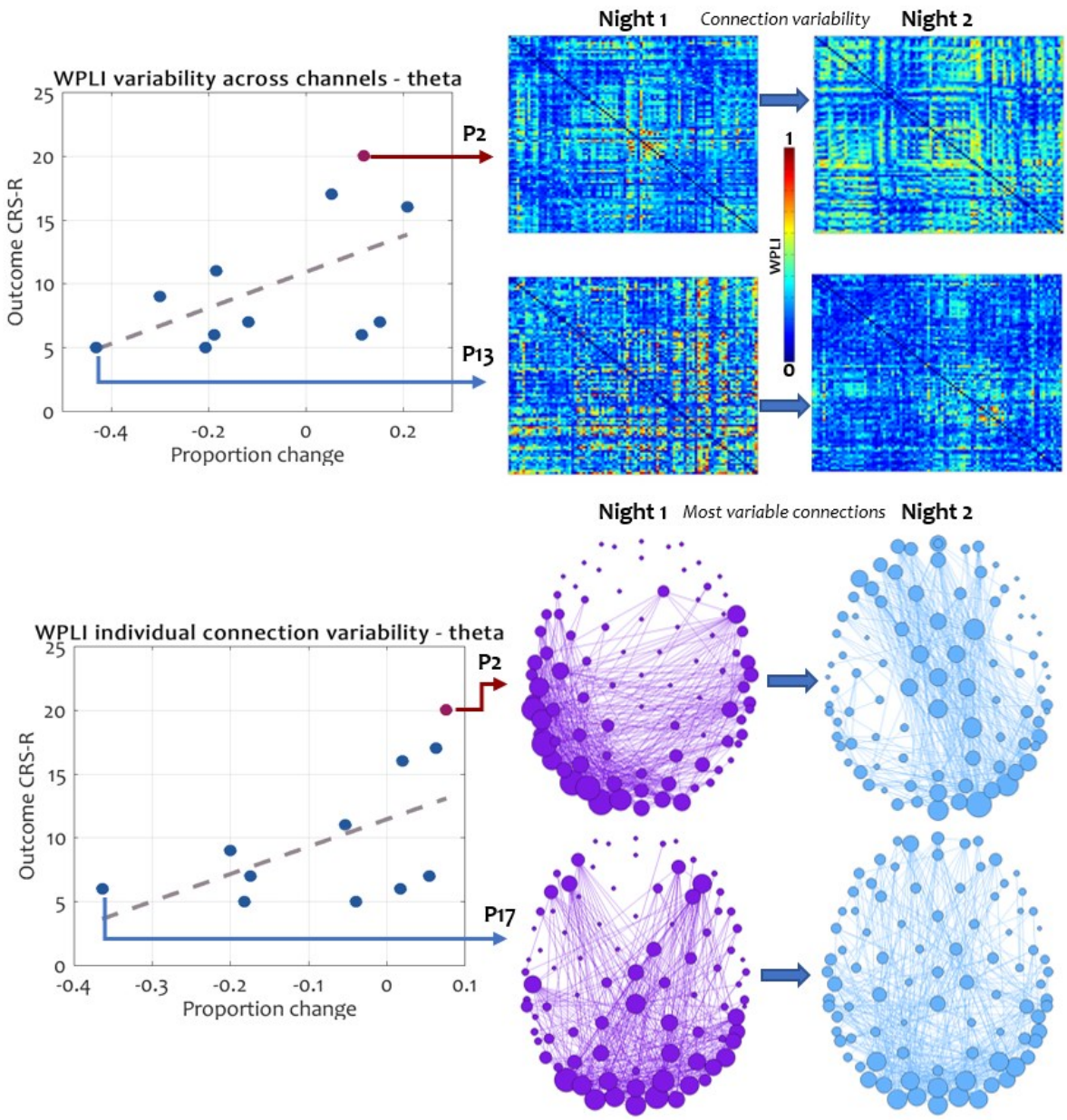

Figure 5.6. ***Theta connectivity variability****. (Top left) Proportion change in momentary WPLI variability of theta networks from first to second session, computed as the median over the standard deviation of the WPLI matrices across time. The dashed line represents a linear fit. (Top right) Illustrations of connectivity variability during the first and second session in two patients, showing the evolution to more variability in P2 (top) and less variability in a patient with poor outcome (bottom). The images represent 91x91 connectivity matrices. (Bottom left) Proportion change in individual connection variability across time from first to second session, computed as the median over the standard deviation of individual connections across time. The dashed line represents a linear fit. (Bottom right) Topographic plots of the strongest 10% values in the matrix of connection variability at each channel pair across time in two patients; the networks in P2 (top) evolve to display most variability in frontoparietal connections during the second session, unlike those of a patient with poor outcome. Node sizes depict its relative number of connections compared to the other nodes in the same network.*

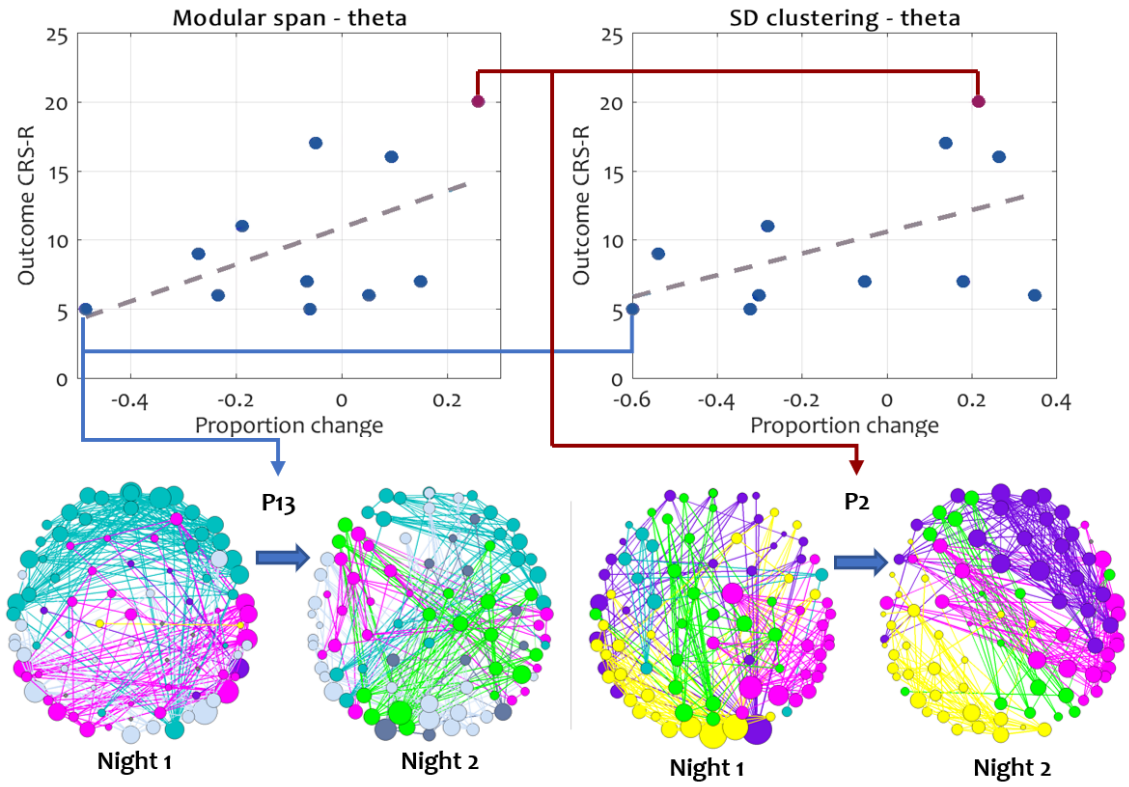

Figure 5.7. ***Theta network modules.*** *Proportion change in theta network median modular span (top left) and median variability of clustering coefficients (top right) over time from first to second session, in all patients. Dashed lines represent linear fits in both panels. The bottom row shows topographical illustrations at a time point where modular span was equal to the median across time. The networks are thresholded at 15%. Different colours show different modules as given by the Louvain algorithm. Only intra-modular connections are plotted. P2 (right) evolves clearly defined modules with long range connections in the second session, while the patient with a poor outcome does not. Node sizes depict the relative number of connections of the node compared to the other nodes in the same network.*

5.3.4. DELTA NETWORKS

The characteristic path length is a proxy for macroscale integration of information across long-range regions in brain networks. Using this measure, we found that increased variability in delta network integration from the first to the second recording after injury was a very good predictor for the outcome of coma across this group of patients, also after correction for multiple comparisons ($p = 0.004$). Node clustering was observed, in particular in patient P2, to vary with the characteristic path length, with a higher path length observed simultaneously with lower clustering (Figure 5.8). Increased variability in node clustering was associated with a better outcome. To some extent, this was also reflected in the standard deviation of small-worldness, where a change towards less variability predicts a poor outcome.

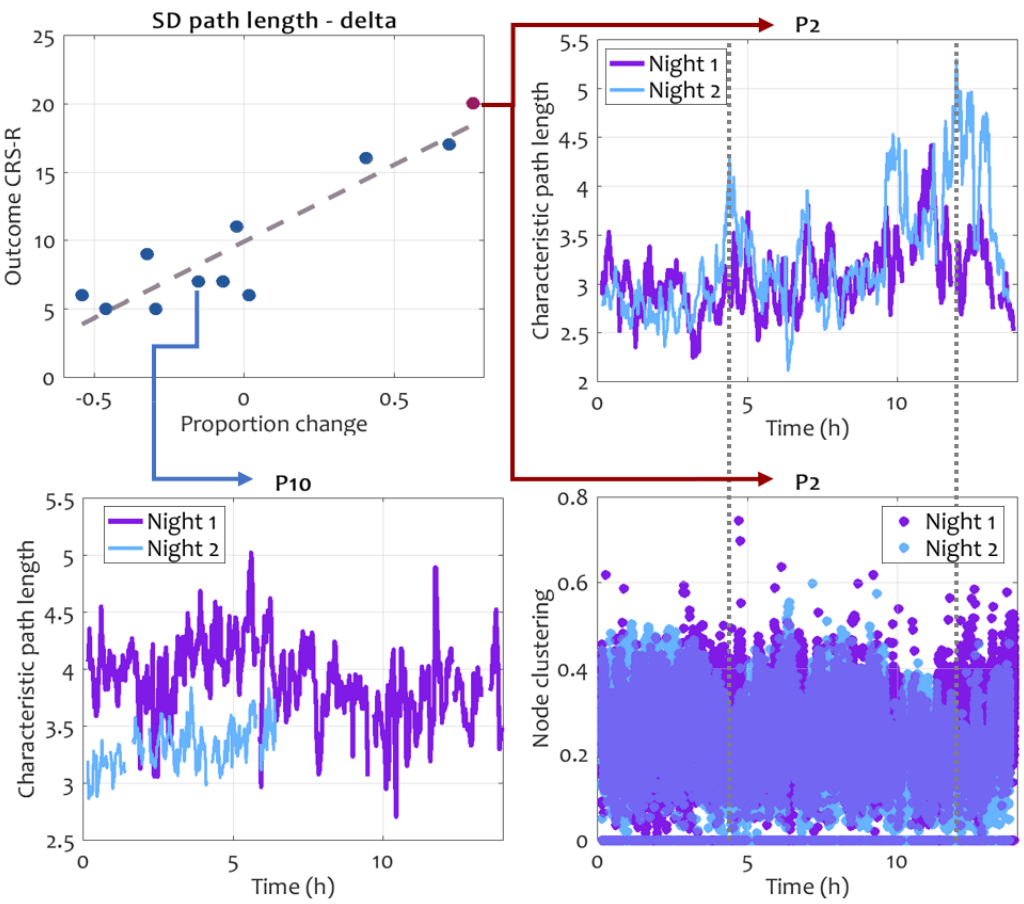

Figure 5.8. ***Delta network characteristic path length variability.*** *(Top left) Proportion change in characteristic path length variability across time in delta networks from first to second session, in all patients. The dashed line represents a linear fit. (Top right and bottom left) Illustrations of timecourses of the characteristic path length in both sessions for P2 and a patient with poor outcome. P2 shows a regular pattern of variation in the second session, whereas the patient with poor outcome does not. Discontinuities in the plots are due to rejected data epochs. (Bottom left) Variability in delta band node clustering in P2. This mirrors the variability in path length, with lower clustering when the path length is higher.*

5.4. DISCUSSION

We proposed a set of computational methods for monitoring and assisting in the diagnosis of comatose patients in the intensive care unit. These methods are inspired by predictions of recent theories of consciousness that a balance between information integration and differentiation in brain networks is essential in maintaining consciousness (Oizumi et al., 2014; Tononi, 2008; Tononi and Edelman, 1998). To achieve this, the WPLI, a measure of connectivity based on the phase lag between two signals (Vinck et al., 2011), was employed to build connectivity networks. This allowed the assessment of the general level of connectivity and its variability in each recording. As a second-order approach, we then used graph theory on the connectivity networks to calculate local and global information processing characteristics, as a proxy for the concepts of differentiation and integration of information.

The crucial observation motivating this study is that covert patterns of reorganised neural activity can reveal a transition towards a higher state of consciousness long before behavioural scores are able to do so. This has been previously observed in chronic disorders of consciousness, where neuroimaging techniques have been successfully applied to find signs of potential covert consciousness in patients displaying no behavioural signs of awareness (Cruse et al., 2011; Fernández-Espejo and Owen, 2013). To assess early neural network changes in comatose patients, we performed two separate overnight recordings in the acute phase, around one and two weeks after the injury, when no behavioural signs of recovery were present in any of the subjects. The change in brain network properties was quantified from the first to the second recording to measure the early progress in the restoration of healthy brain networks. The proportional changes in individual network properties were used as predictors for the eventual outcome of the patients, as assessed behaviourally after two months using the CRS-R score.

We studied the properties of brain networks within three canonical bands: alpha (8-13 Hz), theta (4-8 Hz) and delta (1-4 Hz). The same partitioning has been successfully used in other studies of brain networks in disorders of consciousness (Chennu et al., 2017, 2014). We acknowledge that brain rhythms are severely disrupted in comatose patients with traumatic brain injury; as means to correct for this, the analysis pipeline extracted the maximal connectivity points within the above frequency bands across time windows. Further, to avoid the influence of weak, potentially spurious connections between nodes, the analyses were performed repeatedly over a set of thresholded networks, thus ignoring weak connections. The results across different thresholds were averaged. This approach has been used in other studies (Achard and Bullmore, 2007; Chennu et al., 2014; Lynall et al., 2010). After applying each threshold, graph properties were computed by preserving the weight information in the remaining edges (as opposed to binarizing them), as also advised in previous literature (Monti et al., 2013; Rubinov et al., 2009).

In DoC, key processes that discriminate between vegetative and minimally conscious patients are often found in theta and alpha network activity (Chennu et al., 2017, 2014; Lehembre et al., 2012a; Sitt et al., 2014). Alpha is the dominant rhythm recorded in the human EEG in healthy adults (Klimesch, 2012), therefore an early re-emergence of integrated alpha networks could be a key aspect of a good eventual outcome. Restored alpha network metrics have been recently correlated with good behavioural scores in a large group of DoC patients (Chennu et al., 2017). Indeed, in our study group, patient P2 – the only patient who made a full eventual recovery, despite showing no behavioural signs of improvement at the time of the EEG recordings – showed a unique increase in alpha power from the first to the second recording. This patient also showed an evolution towards a healthy topographical structure in theta and alpha networks akin to those shown by minimally conscious patients at alpha frequencies, with long-range, inter-hemispheric functional modules (Chennu et al.,

2014). Highly restored frontoparietal connectivity was observed in the second recording. Frontoparietal connectivity is known to support a wide range of cognitive abilities that are fundamental to awareness (Bor and Seth, 2012; Laureys and Schiff, 2012; Naghavi and Nyberg, 2005) and is disrupted in impaired states of consciousness (Baars, 2005; De Gennaro et al., 2004; U. Lee et al., 2013). This shows, in agreement with previous studies (Achard et al., 2012; Chennu et al., 2017), that general information integration is not sufficient to produce a good outcome, but the spatial architecture of the recovering functional networks is a key aspect required for the re-emergence of consciousness.

Lower-frequency network properties have also been shown to provide useful markers for diagnostics in impaired consciousness. Stable and increased delta power has been linked to a poor outcome and lack of awareness (Lehembre et al., 2012a; Sitt et al., 2014). On the other hand, an association between variability in delta power and good outcome after coma has also been reported in individual patient cases (Karnaze et al., 1982). A recent study has shown that delta network centrality is a good predictor of outcome in chronic disorders of consciousness (Chennu et al., 2017). Adding to previous results concerning the importance of delta network properties, this chapter provides evidence that increased variability in delta network characteristic path length is a significant predictor of good eventual behavioural outcome. This metric suggests dynamic changes occurring in delta networks during the recovery process. In P2, this was observed as regular periods of stable and variable path length over the course of the night, which might resemble variations that would be expected during sleep (Gross and Gotman, 1999). Importantly, it was not the absolute value of the network metric that was better in P2 compared to other patients (indeed, some patients with poor outcome have a shorter characteristic path length), but the amount of change from the first to second recording, underlining that progress occurring in the acute stage of the injury is crucial for recovery.

Finally, this finding also highlights that variability in network activity is an important predictor of good outcome. The variability of power spectra has been positively associated with better coma outcome (Chiappa and Hill, 1998). In our study, we also found that patients who showed less variability in theta connectivity had a poorer eventual score. Overall, this indicates that the ability of brain networks to change dynamically their regime of long-distance communication is crucial in restoring conscious processes.

The EEG has been used for a long time to assess comatose patients at the bedside. For example, abnormal patterns in the raw EEG (Bagnato et al., 2010; Synek, 1988), evoked potentials (Chiappa and Hill, 1998; Kane et al., 1996) and spectral markers (Kaplan, 2004; Lehembre et al., 2012a, 2012b; Thatcher et al., 1991) have all been shown to be informative in coma prognosis. In comparison to

these approaches, the current study offers a set of EEG tools based on graph theory that are more advanced, but also more indirect. The pipeline involves three levels of approximation: first, the EEG itself is an indirect measure of brain activity, that is able to selectively capture the electric activity of neuronal populations where the neural signal is transmitted perpendicularly to the cortical surface; this orientation may be disrupted due to traumatic injury to the brain. Secondly, the WPLI is a reliable, but still approximate measure of brain connectivity, and a threshold (or a family of thresholds) has to be established to discard edges where it can be assumed there is no significant connection. Finally, graph-theoretical measures may only approximate the real architecture of the brain, at both local and long-range scale. By performing this analysis, one may uncover hidden network patterns that are not visible in the simpler EEG measures mentioned above, but there is also a risk of inaccuracy considering the levels of approximation involved in the computation.

This study provides useful insight into the dynamics of brain networks in acute comatose patients in the early stage after brain injury. However, its limitations should be underlined. Although all patients were diagnosed with traumatic brain injury, the lesion of each individual patient was unique, leading to different structural networks. Within the scope of studying impaired consciousness in a wide range of conditions, we chose to not focus on the clinical details of individual patients, but rather explore whether patterns could be observed despite this heterogeneity. As an advantage of restricting the focus of the study, we created a generalised pipeline that could be rapidly applied at the bedside on any patient in the intensive care unit. This is, of course, complementary to an individual understanding of the structural injury of each patient, which would be crucial in such situations and should be done by clinically-trained personnel. In the context of impaired consciousness, this was an exploratory study that demonstrates how predictions based on theories of consciousness can be applied to EEG recordings in order to help with diagnosis and prognostication in clinical settings.

5.5. ACKNOWLEDGEMENTS

We thank Stein Silva for collecting the EEG data, Srivas Chennu for providing topographic plotting and re-referencing functions, and David Menon for overseeing the clinical care of the patients.

CHAPTER 6

DISCUSSION

This chapter concludes the thesis by recapitulating the main theoretical considerations, methodologies and experimental results presented in the previous chapters, and linking them together in a consistent perspective on the dynamic neural changes that characterise the transitions between states of consciousness. This thesis started by presenting a theoretical perspective on consciousness, followed by an exposition of computational methods suitable for its neuroscientific investigation. This was followed by an experimental investigation of three major transitions of consciousness: sleep, sedation and coma. The results of these investigations are examined in this chapter from the perspective of modern theories of consciousness, which posit that the integration of information across the neural system and the diversity of its neural states are reduced during unconsciousness. This is indeed demonstrated, within certain limitations, across all the three experimental chapters. Nevertheless, more mysteries await for a resolution in this exciting time for the neuroscience and the philosophy of consciousness.

6.1. THEORETICAL OVERVIEW

The thesis began by telling the story of why consciousness is one of the most fascinating topics known to humankind since ancient times. Questions about the subjective perception of the world have appeared pervasively in stories, metaphors, old religions and myths for millennia. The relationship between the mind and the body, famously brought into the spotlight by the writings of René Descartes, has been a long-standing line of enquiry in philosophy. It is only recently that consciousness has become a valid topic for scientific research, as psychology and neuroscience began to understand how brain activity supports cognitive processes that underlie conscious awareness. Conscious experience seems to be built on the foundation provided by the most important cognitive functions of the human brain, such as emotions, memory and attention, but it appears to be more than a sum of these functions. As an independent concept in cognitive research, consciousness deserves a proper definition – which it does not yet have across all disciplines.

Section 1.3 highlighted several theoretical milestones that have helped with crystallising the notion of consciousness from the perspective of both philosophy and neuroscience. There are two

complementary aspects of consciousness that are apparent in such works. One is the first-person, subjective experience of being conscious, either as a state (being awake) or in the sense of being conscious *of* something. The latter, in particular, is closely linked with the concept of qualia: the ineffable properties of reality that we introspectively have conscious access to, such as the redness of an apple or the sweetness of chocolate. In simple terms, it is what causes Mary the neuroscientist, who learned everything about colour vision while living in a black-and-white room all of her life, to learn something new when she actually sees the colour red for the first time. The hard problem of consciousness asks why this experience exists in the first place as a companion to neuronal firing. While intuitively useful, there is a problem with defining consciousness in this manner: the first-person, subjective conscious experience is not directly measurable and hence cannot, by itself, form the object of scientific investigation.

The second approach is the third-person, objective method of studying consciousness. We can take a small leap of faith to believe that verbal reports and behavioural responses from other people are a good enough indicator of conscious experience, as opposed to being traits of so-called philosophical zombies. They are not perfect measures, as emphasised in section 1.5: it is possible to be conscious without the possibility to respond, and unconscious while producing valid responses. However, reports and responsiveness provide an acceptable objective proxy for the first-person subjective experience. In this way, we can focus on unravelling the architecture of our conscious space without worrying about why it exists, while also not ignoring its highly subjective dimension. A full consideration of such neurophenomenology is currently the real problem that awaits to be gradually resolved by neuroscience.

The idea of neural correlates of consciousness is relatively recent (Chalmers, 2000; Crick and Koch, 1990; Metzinger, 2000). The neural correlates of the contents of consciousness, which refer to the qualities of the world that we are aware of, have been studied and debated extensively (Boly et al., 2017; Odegaard et al., 2017). There is, however, another aspect that deserves attention: the transitions between levels (or states) of consciousness. Here, we are not interested in *what* we are aware of at a given moment, but how *awake* and *responsive* to the external world we are. In other words, the aim is to investigate how consciousness is lost and regained. This not only a fundamental question for basic research, but also a critical question in clinical settings involving patients recovering from brain injury or monitoring anaesthesia during surgery. As we saw in section 1.4, unconsciousness can occur due to natural, pharmacological and pathological causes: sleep, sedation and anaesthesia, disorders of consciousness and coma. Each of these three consciousness transitions is devoted a full experimental chapter (3, 4 and 5) of the thesis.

The first chapter concluded with the introduction in section 1.6 of several theories of consciousness that can frame the computational investigations of these topics. In particular, the integrated information theory (IIT) is especially attractive, as it aims to bridge the gap between the subjective reality of conscious experience and objective reports by creating postulates about physical rules that must be implemented by neural systems, inspired by the phenomenological structure of the world. Despite the fact that its axioms and the derivation of its postulates can be challenged (Bayne, 2018), the advantage of IIT consists of its top-down approach anchored in phenomenology, which is in contrast with theories that merely aim to correlate electrophysiological signatures with subjective reports. Moreover, IIT and other modern theories of consciousness do not restrict their scope to finding neural correlates within individual brain areas, but instead examine high-level patterns of connectivity and functioning across the brain as a system.

The results of this thesis are generally in agreement with the three theories of consciousness presented in the introduction: the global workspace theory, the dynamic core hypothesis and the integrated information theory. However, the scope of the evidence presented here towards all of these theories is limited to levels, as opposed to contents, of consciousness. This is important in particular for IIT, whose theoretical development is centred around phenomenal experience. Furthermore, the mentioned theories make predictions that qualitatively refer to the existence or non-existence of functional connections between brain areas, or to properties of local activity, such as complexity (the diversity of possible states); they do not make predictions regarding particular frequencies at which information would be generated or exchanged. For this reason, the evidence provided by spectral analyses should be interpreted with caution. This being said, we may assume that some frequencies, such as alpha, are fundamental in information processing across brain networks in healthy, responsive adults, as discussed further in the next section.

According to the global workspace theory, a lower level of consciousness should be associated with a disconnection across the neural system, and thus less ability of information exchange across specialised submodules. This agrees with the sleep experiment results, where alpha networks break down when the participants are unresponsive. It also agrees with the sedation experiment results, where there is less mutual information sharing at moderate sedation at temporal separation parameters corresponding to frequency bands up to alpha. Finally, the coma experiment results can be interpreted as supportive of this idea, as the variability of the path length can suggest a recovery in the dynamics of global information exchange, while the restored long-range alpha and theta connectivity of the patient with the best outcome supports the same prediction; however, not all measures of information exchange across brain network modules were found to be predictive of

eventual outcome, and caution should be used in this interpretation due to the high heterogeneity of coma cases present in this group.

From the perspective of IIT and the dynamic core hypothesis, neural complexity and information integration should be decreased in lower levels of consciousness. If we assume alpha networks to be fundamental for cognition, the breakdown of alpha connectivity during unresponsiveness is supportive of this prediction, along with the decrease in mutual information index during sedation. In coma, the same interpretation regarding the variability of delta band path length and alpha long-range connectivity that supports the global workspace theory is in agreement with mutual information exchange. Furthermore, it is shown that propofol sedation is indeed associated with lower Lempel-Ziv complexity of the EEG, as predicted by the two theories. However, a small variation in the computation, which affects whether the complexity reflects more prominently the temporal or the spatial patterns of neural activity, led to a different result in the group that was still responsive under moderate sedation: temporal complexity is higher, whereas spatial complexity is lower. The Lempel-Ziv complexity of the EEG is an estimative measure, so this difference between the two variants of the measure does not directly speak for or against any direct prediction of either IIT or the dynamic core hypothesis. Conversely, a possible interpretation that was discussed here is that spatial complexity tracks the level of drug, whereas the temporal complexity tracks the level of consciousness. These results invite further investigation into complexity measures and their relation to levels of consciousness.

To sum up overall, the results of this thesis are generally in agreement with the three theories discussed, but only within the limitations emphasised above.

6.2. FROM PHENOMENOLOGY TO COMPUTATIONAL TOOLS

When searching for neural signatures that underlie distinct conscious states, there are two key properties predicted by modern theories to be essential in sustaining consciousness: information integration and differentiation. Phenomenologically, integration refers to the unity of conscious experience: neither can it be decomposed into parts, nor can we have two separate conscious experiences at the same time. Differentiation is related to entropy and refers to the highly informative character of every single experience: our neural architecture is capable of representing an unimaginably vast number of other conscious experiences comprising a combination of different colours, sounds, objects, people, spatial environment, emotional state and so on – all at the same time.

Computationally, measures like ϕ, which was proposed by IIT to compute the causality generated by a system and which would directly address the degree of integration and differentiation at neuronal level, are currently computationally intractable to estimate for the human brain. This is because it is currently unknown how to describe the brain as a system in the form required for computing this measure. Instead, computational tools that quantify these properties using surrogate approaches have been proposed (Tegmark, 2016). In practice, to compute such measures of consciousness, several established methodologies of analysing brain activity are often used.

Chapter 2 investigated three such methodologies and their significance in consciousness research along with other traditional methods of EEG assessment: spectral measures (power and phase-lag connectivity), information-theoretical measures (signal complexity and mutual information sharing) and fast-paced global activity dynamics (EEG microstates). The EEG was chosen for the experimental investigations of this thesis due to several of its advantages over other neural recording methods. It is fast and easy to use in any setting, including a clinical environment, rendering it suitable for monitoring anaesthesia during surgery or diagnosing patients with brain injury in intensive care. It also provides very high temporal resolution. All experiments in this study employ at least 62 electrode channels, which also offers a good resolution in sensor space. For this thesis, it was decided to keep all analyses in sensor space, to limit the assumptions and parameters entailed by the analytical approach. Furthermore, we were less focused on anatomical localisation than on characterising the functionally relevant dynamics observable in the EEG. Moreover, in the clinical cohort of comatose patients with traumatic brain injury analysed in Chapter 5, an analysis pipeline for convenient use at the bedside of any patient would not have been generalisable in source space, given the unique aetiology of every injury.

As described in section 2.2.1, one of the oldest methods of analysing the EEG is investigating its spectral power. This methodology has been developed and validated in a data-driven manner over many decades. Several rhythms have been ubiquitously found in the human EEG in association with distinct states of consciousness, such as the alpha rhythm observable during relaxed wakefulness. During unconsciousness, alpha waves disappear, and lower-frequency power emerges. The frequency bands observed to covary with the state of consciousness (alpha, theta and delta) have served as canonical bands in all chapters of this thesis. The fixed choice of frequency bands has both advantages and disadvantages. In healthy adults, although a high variability in the peak frequency of individual rhythms has been reported (Klimesch, 1999), the frequency bands established here have been validated across almost a century of studies, and therefore we do not expect individual variability to be a particular issue in the sleep and sedation analyses. On the other hand, this is a more delicate topic in the group of acute comatose patients. These patients rarely show meaningful power at higher

frequencies, and canonical bands may restrict the observability of changes occurring within narrow bands at low frequencies (Lehembre et al., 2012a). However, by tracking the peak in power and connectivity in each band, it was ensured that the most prominent component within each band was tracked. Moreover, in Chapter 5, a patient whose alpha networks re-emerged while no behavioural improvement was yet visible, but who had an excellent eventual outcome, demonstrated that it is important to monitor canonical frequency bands. This also allowed the development of a general pipeline that could find utility in the clinical context if developed further.

Spectral power measures activity within individual neural modules, but it is widely thought that such modules communicate across long-range distances in order to give rise to coherent behaviour, as opposed to individual phenomenological entities being coded by single specialised assemblies of cells (Singer and Gray, 1995). This is in line with the integrative property of conscious experience as posited by theories of consciousness like IIT. The hallmark of long-distance communication between neural assemblies is considered to be phase synchronisation (Fell and Axmacher, 2011; Sauseng and Klimesch, 2008; Varela et al., 2001). Section 2.2.2 justified why the WPLI represents a good measure for scalp-level spectral connectivity, for reasons including its relative robustness to volume conduction. The WPLI was then used in Chapters 3 and 5 of this thesis, as well as in a previous study on the same dataset employed in Chapter 4 (Chennu et al., 2016a).

In the landscape of spectral measures, this thesis confirms that alpha power and networks are particularly important for sustaining a healthy and responsive state of consciousness. Historically, the alpha rhythm was the first prominent feature observed in the EEG (Berger, 1930). Its strongest sources appear to be in the occipital-parietal area, but alpha oscillations can be recorded from the whole scalp, including frontal areas (Nunez et al., 2001). Because the alpha rhythm is especially prominent during relaxation with closed eyes and is suppressed by eye opening, motor movements, challenging mental tasks, or unconsciousness, it has been suggested that it indicates an idle state of the wakeful brain (Pfurtscheller et al., 1996). More recent findings that visual task performance depends on the phase of the alpha rhythm at stimulus presentation suggest that alpha oscillations could represent periodic inhibitory waves (Busch et al., 2009; Mathewson et al., 2009). However, there is also evidence that alpha networks could play an important role in executive and attentional task-relevant processing (Palva and Palva, 2011). While the role of the alpha rhythm is still being uncovered, its importance as a marker of healthy wakefulness is established in existing literature and confirmed in this thesis.

In addition to spectral tools, information-theoretical measures, presented in section 2.3, can be an appropriate tool of investigation, particularly within the framework of consciousness theories that make predictions about information integration and complexity. These measures follow the idea that

information in general (in the entropic sense proposed by Claude Shannon) underlies the conscious experience – not necessarily in spectral form. The two measures for information integration and for information diversity proposed in Chapter 4 of this thesis are the wSMI and the LZ complexity. The former is based on pattern-matching, while the latter is derived from a popular data compression algorithm used in computer science. How are these related to their corresponding spectral measures used in Chapters 3 and 5 – connectivity and power? The wSMI is known to capture synchronisation bounded by specific frequencies, depending on the parameters employed in its calculation (King et al., 2013). At the same time, as described in Chapter 4, LZ complexity is not fully explained by power changes, a finding discussed in other studies employing the same measure (Schartner et al., 2017a, 2017b, 2015). Overall, there seems to be a partial degree of overlap between spectral and information-theoretical measures, but more research is needed to establish an exact relationship. Finding new and better metrics for assessing information integration in complex systems constitutes an active area of research (Barrett and Seth, 2011; Oizumi et al., 2016; Tegmark, 2016).

Finally, in section 2.5, one more methodological tool was added to the toolkit for exploring consciousness levels. Recent literature has underlined the importance of fast-paced millisecond-level dynamics of brain networks (Baker et al., 2014; Vidaurre et al., 2016). According to the theoretical framework, the diversity of states that the neural system can express changes with the level of consciousness; therefore, we can also expect alterations in the fast temporal dynamics of global brain activity. The EEG can be fragmented into a sequence of a consistent set of topographical maps that last for tens of milliseconds each. The sequence of microstates has been previously hypothesized to be linked to cognitive processing (Lehmann, 1990). As a data-driven method, the method of electric microstates was used in conjunction with spectral analysis in Chapter 3, to reveal the transitory dynamics of brain networks during the onset of sleep.

6.3. LESSONS FROM SLEEP, SEDATION AND COMA

Armed with a theoretical framework on neural integration and differentiation and a set of methodological tools to quantify these properties, the next chapters investigated three situations where a transition between states of consciousness and unconsciousness occurs: first, the loss of responsiveness in the most familiar of all transitions: sleep; secondly, pharmacologically-induced unconsciousness: sedation; and finally, the pathological loss of consciousness due to traumatic brain injury: coma. The novel findings regarding each of these transitions are summarised in this section.

Chapter 3 investigated the neural dynamics of falling asleep in a group of healthy subjects performing an auditory discrimination task. First, it was found that alpha power and frontoparietal connectivity

were significantly higher while they were responsive, whereas theta power was higher when they became unresponsive. Next, it was found that microstate dynamics slow down after responsiveness is lost due to drowsiness, and a unique microstate (D) was identified whose increased duration predicted behavioural unresponsiveness on a trial-by-trial basis. By combining information about spectral brain connectivity and electric microstates, it was revealed that microstate D also uniquely captures a specific increase in frontoparietal theta connectivity, a putative marker of the loss of consciousness prior to sleep onset. This finding also highlights transient and distinct brain networks active during the onset of sleep.

Using a similar paradigm with respect to the loss of responsiveness to auditory stimuli, Chapter 4 examined a group of healthy subjects undergoing mild and moderate sedation with propofol. The LZ complexity was used to assess differentiation in the neural signal and the wSMI was used to assess information integration across brain networks. An intriguing dissociation was found between responsiveness and drug level in blood during sedation: responsiveness is best predicted by the temporal complexity of the signal at single channels and by information integration in theta and delta bands, whereas drug level is best predicted by the complexity of spatial electric activity patterns and information integration in alpha bands. Distinct signatures of drug level in blood and the level of responsiveness in sedation have been previously reported (Barttfeld et al., 2015; Chennu et al., 2016a). In section 4.4.3, it was emphasised how the same dissociation between temporal and spatial complexity found in this thesis (namely, that spatial complexity decreases in monotonically, whereas temporal complexity increases in sedated subjects who still respond, before decreasing after they stop responding) is also evident in a recent study of LZ complexity in propofol anaesthesia (Schartner et al., 2015). A similar evolution of the two types of LZ complexity was described in an entirely different alteration of consciousness involving psychedelic substances (Schartner et al., 2017a), which suggests that moderate doses of propofol could cause an altered state of consciousness while the subject is still responsive. Indeed, this agrees with behavioural reports of altered subjective experiences during propofol sedation (Balasubramaniam and Park, 2003; Brandner et al., 1997).

Finally, Chapter 5 investigated brain connectivity in the overnight EEG recordings of a group of acute comatose patients, with the aim applying knowledge gained from the underpinning neuroscience research employed in the previous chapters to create a computational pipeline for assessing brain dynamics that could be feasibly deployed in the clinical context, at the bedside of the patient. This was motivated by an existing need for clinical tools that could assist with diagnostics and prognosis after brain injury, both in acute and chronic cases. Currently, prognostication in coma is difficult (Stevens and Sutter, 2013) and the level of consciousness in patients is prone to misdiagnosis (Schnakers et al., 2009). Towards building methods to address this problem, graph theory was applied on alpha, theta

and delta networks obtained using the WPLI on two nights during the acute phase of recovery in a small group of patients in the intensive care unit. The proportional change in network properties from the first to the second night served as a measure of early dynamic reorganisation of brain networks in the healing process. This change was correlated with behavioural improvement measured after two months, representing the long-term outcome of the patient. Amongst other trends, it was found that increased variability in the characteristic path length of delta networks early after injury predicted the eventual coma recovery score. This suggested that variability in low-frequency network characteristics could be important in the early neural reorganisation after injury. A remarkable case study of an individual patient was also considered in depth, in whom the early re-emergence of frontoparietal alpha connectivity predicted a full recovery long before any behavioural improvement occurred. This confirmed that the neural networks supporting cognition and consciousness can re-emerge in disorders of consciousness even when the patient is not able to overtly express this, an important finding that has reshaped the public and clinical view of such disorders over the last decade since the discovery of covert cognition in the vegetative state (Owen et al., 2006). Due to the clinical nature of this dataset, the relatively modest sample size and the considerable heterogeneity of each patient's individual injury, the study presented in this chapter had limitations that warranted careful interpretation of the results, as discussed in section 5.4.

6.4. BEHAVIOURAL CONSIDERATIONS

Chapters 3 and 4 investigated the transition to unconsciousness during drowsiness and sedation using similar behavioural measures: responsiveness to simple auditory tasks. In comparing these experiments, one concern could be that the difficulty of the two tasks was different. In the semantic classification task employed in the sleep study, participants were required to classify the stimulus word as an object or an animal. In contrast, in the perceptual task used in the sedation study, participants had to discriminate a buzz from a noise sound. Although the latter task can appear to be easier, it can be argued that it was not the difficulty of the task that stopped the participants from responding, but the transition into the state of unconsciousness itself. A previous study on the same sleep dataset employed in this thesis (Kouider et al., 2014) found that participants were still producing a lateralised readiness potential corresponding to the correct response even when they failed to produce a motor response. This indicates that cognitive processes performing a non-trivial task can still be active after the loss of responsiveness, so it must be a different dynamical change in the state of the brain that prevents the motor response from occurring. Therefore, the points of loss of responsiveness in the two experiments can be considered to be comparable. Finally, it was further discussed in section 1.5, but also across Chapters 3 and 4, how responsiveness does not always equate

consciousness, but rather provides a proxy to explore an essential part of the transition from consciousness to unconsciousness in healthy subjects.

On the other hand, in Chapter 5, behavioural scores of responsiveness in patients reflect a completely different dimension of consciousness. Here, the CRS-R scale is used as an outcome measure for the comatose patients. This assesses their sensory awareness and motor ability to produce a combined estimation of the level of recovery. The following assumption is therefore made: that behavioural responsiveness reflects well enough the true level of impairment after a stabilisation period (two months), which is not evident in the acute phase of the injury due to the immediate impact of the traumatic event. Indeed, the predictive value of the CRS-R score has been demonstrated after a mean of 48 days after injury (Bodien et al., 2016), but prognostication using behavioural or other types of markers is still challenging in the acute phase of the injury (Stevens and Sutter, 2013). Here, the EEG analyses in the acute phase of the injury seek to uncover patterns of re-emerging neural activity that are not yet visible behaviourally. The case study of patient P2 that was discussed in this chapter, as well as other notable studies (Cruse et al., 2011; Fellinger et al., 2011; Harrison and Connolly, 2013; Owen et al., 2006), demonstrate that behaviour is indeed sometimes preceded by covert changes in neural activity in disorders of consciousness. Does that mean patients with covert cognitive function are conscious? This is a matter of debate (Overgaard, 2009; Overgaard and Overgaard, 2011), but at least it can certainly be argued that the restoration of such processes, even in covert form, is required for reaching a healthy state of consciousness again.

6.5. DYNAMICAL DIVERSITY OF NEURAL STATES

Theories of consciousness predict a decline in neural activity differentiation in conjunction with the loss of consciousness. In agreement with this prediction, all experiments in this thesis have shown that a loss in the dynamic range of states of the neural system corresponds to the transition towards unconsciousness.

In Chapter 3, it was shown that the changing rate of electric microstates of the brain significantly slows down when subjects stop responding as they fall asleep. In this analysis, there is no loss in the number of states, represented by the four microstate topographies, but there is a reduction in the temporal variability of their dynamics. Although research on the functional (Milz et al., 2015; Seitzman et al., 2016) and structural (Britz et al., 2010; Pascual-Marqui et al., 2014) meaning of the four microstates currently shows no clear consensus on their neural origin, the EEG microstates nonetheless reflect a global state of the brain which, in line with theoretical predictions, becomes less temporally diverse during the onset of sleep.

Next, in Chapter 4, a reduction in diversity as measured directly from the complexity of the EEG signal was also observed to occur with propofol sedation. Intriguingly, however, two flavours of complexity, which measure temporal and spatial complexity, respectively, were found to track different aspects of sedation. Temporal complexity tracked the responsiveness of the subjects, whereas spatial complexity tracked their blood concentration of propofol. Although there was a high correlation between these two measures, their distinct evolution with increasing propofol dosage was statistically significant. The functional meaning of these distinct signatures was discussed in section 4.4.3, as well as in other studies that found different brain signatures of drug level and responsiveness (Barttfeld et al., 2015; Chennu et al., 2016a).

The reduction in the diversity of network states predicted by IIT during unconsciousness could occur both spatially, with reduced integration across the system leading to a smaller amount of simultaneous interactions (with distinct states being more similar), and temporally, with decreased information exchange leading to a more restricted number of states that the system can be in. In the sedation experiment, responsiveness, which is a more faithful indicator of the level of consciousness compared to the drug concentration in blood, was found to be best tracked by temporal complexity. Further, the slowing down of microstate dynamics also indicates a decline in temporal complexity. Therefore, a finding bridging across results in this thesis is that a temporal reduction in neural complexity is a valuable marker of the transition towards unconsciousness.

The findings presented in Chapter 5 about the group of comatose patients can also be interpreted in a manner that underlines the importance of temporal diversity of neural states. Amongst all graph-theoretical measures tested as predictors for the eventual patient outcome, the variability of path length in delta networks was the only statistically significant predictor. This indicates that an increased number of global configurations over time in the delta networks, as opposed to a change in the properties of the states themselves, is beneficial in the restoration of consciousness.

6.6. LONG-RANGE INTEGRATION IN BRAIN NETWORKS

It is further predicted by modern theories of consciousness that integration is fundamental in brain networks sustaining consciousness. In congruence with this prediction, it was demonstrated across this thesis how integration across brain networks supports healthy wakefulness. In particular, scalp-level frontoparietal connectivity was shown to be an essential signature of consciousness.

In Chapter 3, it was found that alpha frontoparietal connectivity was significantly higher while participants were still responsive compared to the periods when they lost responsiveness as they fell asleep. This validates previous results on the loss of alpha connectivity in both sleep and other

consciousness transitions, such as anaesthesia (Chennu et al., 2016a, 2014; Ogilvie, 2001; Tanaka et al., 2000, 1998; Wright et al., 1995). In contrast, theta connectivity increased during unresponsiveness, although not significantly. A reduction in frontoparietal activity during unconsciousness has also been reported using other imaging modalities (Kajimura et al., 1999; Larson-Prior et al., 2011; Spoormaker et al., 2012).

Chapter 4 employed a dataset where a previous study (Chennu et al., 2016a) identified that frontoparietal connectivity is preserved during sedation only in the subjects who remained responsive to the auditory task. In this thesis, connectivity was investigated using the symbolic mutual information, a measure originally proposed for classifying the clinical state of patients with disorders of consciousness (King et al., 2013). The transition to unconsciousness was accompanied by lower wSMI in delta, theta and alpha bands, and higher wSMI in beta band, but no clearly topographical patterns of connectivity were associated with the descent to unconsciousness. Overall, a note of caution is appropriate here, clarifying that this measure was considerably variable across and within subjects. Hence, further validation is required from future studies to assess its efficacy to capture and track brain connectivity.

Finally, in Chapter 5, it was found that increased variability in the characteristic path length of delta networks early after traumatic injury predicts a positive outcome in comatose patients. The characteristic path length is a measure of long-range integration across brain networks. Interestingly, it was not an increased path length in itself that predicted a good outcome, but its variability in overnight recordings. This suggests that a dynamic repertoire of global network configurations with respect to information integration is required for the recovery of healthy wakefulness after coma. Furthermore, a case study was presented where a severely brain-injured patient with a nevertheless positive eventual outcome displayed a remarkable early re-emergence of alpha frontoparietal connectivity while still being behaviourally unresponsive during the acute phase of coma.

Connectivity between frontal and parietal brain areas has been linked to a wide range of cognitive functions (Babiloni et al., 2004; He et al., 2007; Naghavi and Nyberg, 2005) and its suppression has been demonstrated in sleep, anaesthesia and disorders of unconsciousness (Boly et al., 2012; Chennu et al., 2016a, 2014; Kajimura et al., 1999; Laureys and Schiff, 2012; Lee et al., 2009a; Spoormaker et al., 2012), which is in agreement with the hypotheses and results in this thesis. However, it is debated whether frontoparietal connectivity represents, in itself, a true neural correlate of consciousness, or it is rather confounded by the recruitment of processes intimately associated with conscious processing, such as attention, memory and executive control. For example, recent evidence indicates that frontoparietal activity is not directly elicited by consciously attending stimuli of relevance in a

detection task, but is instead activated during goal-related tasks (Farooqui and Manly, 2017). Moreover, no-report paradigms support the idea that frontal activity is not required for conscious perception (Tsuchiya et al., 2015). On the other hand, it can be argued that cognitive subsystems that sustain the full conscious experience, as opposed to raw conscious perception, include functions like memory, attention and biologically-driven goals, therefore rendering the frontoparietal network a necessary part of the comprehensive architecture required for consciousness (Bor and Seth, 2012). This topic can further be contextualised in the ongoing search for the location of the neural correlates of contents of consciousness. There is currently lively debate on whether the posterior (Boly et al., 2017) or the frontal (Odegaard et al., 2017) areas of the cortex are required for awareness of specific perceptual experiences.

The results of this work offer a limited degree of evidence towards the hypothesis that frontoparietal networks, especially at higher frequencies, are very important in supporting a level of consciousness where responsiveness is observed. This does not offer evidence regarding whether the anterior or the posterior areas of the brain are required for consciousness, but rather suggests that the communication between these areas is essential. More specifically, in sleep, frontoparietal alpha connectivity, which is a marker of healthy wakefulness, is significantly weaker when responsiveness is lost; however, theta frontoparietal connectivity is stronger. In sedation, mutual information sharing is lower at a temporal separation corresponding to delta, theta and alpha frequency bands; however, it is higher for gamma band. In acute coma, alpha, theta and delta network reconstruction is associated with a better eventual outcome; however, this is reflected in only some of all the graph measures that were tested. Although overall these results suggest that changes in frontoparietal connectivity are associated with a change in level of consciousness, this relationship is not straightforward. Moreover, most electroencephalography measures employed in this thesis cannot answer the question of whether there is activation at a particular time in a particular brain region, but rather approximate how much activity or synchronicity there is when measured within certain parameters, such as a particular frequency band. There is, however, one result presented in his thesis that may speak to this debate, which is that complexity is significantly higher in frontal areas in sedated but responsive subjects, as compared to unresponsive subjects. This could suggest that frontal areas are indeed important in sustaining consciousness. Finally, these results are not directly informative for any debate regarding the contents, but only the levels, of consciousness.

The debates regarding the role of the front of the brain, the back of the brain and their connectivity patterns will likely take a long time to be resolved, experimentally and conceptually, but it is important to keep an open mind and respect the diversity of perspectives currently co-existing in this research field.

6.7. CONCLUSIONS

This thesis has taken the reader through a journey around one of the most fascinating topics ever known to humankind: consciousness. It started broadly, from the historical roots of this concept, and then focused on a specific goal: understanding transitions between levels of consciousness. Equipped with a modern theoretical framework and a set of computational methods appropriate for the neuroscientific investigation of consciousness, three distinct such transitions were investigated: sleep, sedation and coma. Each of these transitions provides a unique window into the loss of consciousness, yet a number of similarities were identified by an examination from a common theoretical perspective. Most importantly, this thesis has demonstrated that there is something quantifiable that is lost in brain activity as consciousness fades: the *dynamical diversity* of neural signals, as well as the *long-range integration* across brain networks, especially alpha band connectivity. The value of these results is two-fold. On the one hand, such advances can find practical applications in the clinical context, where there is a need for better tools to diagnose, evaluate and track patients undergoing anaesthesia or recovering from brain injury. On the other hand, these results contribute to the theoretical advancement of theories of consciousness by offering a unique perspective on its transitions.

REFERENCES

Abásolo, D., Simons, S., Morgado da Silva, R., Tononi, G., Vyazovskiy, V. V., Abasolo, D., Simons, S., Silva, R.M. da, Tononi, G., Vyazovskiy, V. V., 2015. Lempel-Ziv complexity of cortical activity during sleep and waking in rats. J. Neurophysiol. 113, 2742–2752.

Achard, S., Bullmore, E., 2007. Efficiency and cost of economical brain functional networks. PLoS Comput. Biol. 3, 0174–0183.

Achard, S., Delon-Martin, C., Vértes, P.E., Renard, F., Schenck, M., Schneider, F., Heinrich, C., Kremer, S., Bullmore, E.T., 2012. Hubs of brain functional networks are radically reorganized in comatose patients. Proc. Natl. Acad. Sci. U. S. A. 109, 20608–13.

Alkire, M.T., Hudetz, A.G., Tononi, G., 2008. Consciousness and Anesthesia. Science (80-.). 322, 876–880.

Allen, N.J., Barres, B.A., 2009. Neuroscience: Glia - more than just brain glue. Nature 457, 675–677.

Amantini, A., Grippo, A., Fossi, S., Cesaretti, C., Piccioli, A., Peris, A., Ragazzoni, A., Pinto, F., 2005. Prediction of "awakening" and outcome in prolonged acute coma from severe traumatic brain injury: Evidence for validity of short latency SEPs. Clin. Neurophysiol. 116, 229–235.

Andrillon, T., Poulsen, A.T., Hansen, L.K., Léger, D., Kouider, S., 2016. Neural Markers of Responsiveness to the Environment in Human Sleep. J. Neurosci. 36, 6583–6596.

Araújo, A.M., Machado, H.S., Falcão, A.C., Soares-da-Silva, P., 2017. Reliability of body-weight scalars on the assessment of propofol induction dose in obese patients. Acta Anaesthesiol. Scand. 1–10.

Arbour, C., Baril, A.-A., Westwick, H.J., Potvin, M.-J., Gilbert, D., Giguère, J.-F., Lavigne, G.J., Desautels, A., Bernard, F., Laureys, S., Gosselin, N., 2016. Visual Fixation in the ICU: A Strong Predictor of Long-Term Recovery After Moderate-to-Severe Traumatic Brain Injury. Crit. Care Med. 44, 1.

Århem, P., Liljenström, H., 1997. On the Coevolution of Cognition and Consciousness. J. Theor. Biol. 187, 601–612.

Aristotle, 350BC. De Anima.

Baars, B.J., 2005. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. Prog. Brain Res. 150, 45–53.

Baars, B.J., 1988. A Cognitive Theory of Consciousness. Cambridge University Press.

Babiloni, C., Carducci, F., Vecchio, F., Rossi, S., Babiloni, F., Cincotti, F., Cola, B., Miniussi, C., Rossini, P.M., 2004. Functional frontoparietal connectivity during short-term memory as revealed by high-resolution EEG coherence analysis. Behav. Neurosci. 118, 687–697.

Bachman, P., Sordoni, A., Trischler, A., 2016. Towards Information-Seeking Agents 1–11.

Badia, P., Wright, K.P.J., Wauquier, A., 1994. Fluctuations in single-hertz EEG activity during the transition to sleep. Sleep onset Norm. Abnorm. Process. 201–218.

Bagnato, S., Boccagni, C., Prestandrea, C., Sant'Angelo, A., Castiglione, A., Galardi, G., 2010. Prognostic value of standard EEG in traumatic and non-traumatic disorders of consciousness following coma. Clin. Neurophysiol. 121, 274–280.

Bai, D., Pennefather, P.S., MacDonald, J.F., Orser, B.A., 1999. The General Anesthetic Propofol Slows Deactivation and Desensitization of GABA A Receptors. J. Neurosci. 19, 10635–10646.

Bai, Y., Liang, Z., Li, X., Voss, L.J., Sleigh, J.W., 2015. Permutation Lempel–Ziv complexity measure of electroencephalogram in GABAergic anaesthetics. Physiol. Meas. 36, 2483–2501.

Baker, A.P., Brookes, M.J., Rezek, I.A., Smith, S.M., Behrens, T., Probert Smith, P.J., Woolrich, M., 2014. Fast transient networks in spontaneous human brain activity. Elife 3, 1–18.

Balasubramaniam, B., Park, G.R., 2003. Sexual hallucinations during and after sedation and anaesthesia. Anaesthesia 58, 549–553.

Bareham, C.A., Manly, T., Pustovaya, O. V, Scott, S.K., Bekinschtein, T.A., 2014. Losing the left side of the world: rightward shift in human spatial attention with sleep onset. Sci. Rep. 4, 5092.

Barker, W., Burgwin, S., 1949. Brain wave patterns during hypnosis, hypnotic sleep and normal sleep. Arch. Neurol. Psychiatry 62, 412–20.

Barrett, A.B., Seth, A.K., 2011. Practical Measures of Integrated Information for Time-Series Data. PLoS Comput. Biol. 7, e1001052.

Barry, R.J., Clarke, A.R., Johnstone, S.J., Magee, C.A., Rushby, J.A., 2007. EEG differences between eyes-closed and eyes-open resting conditions. Clin. Neurophysiol. 118, 2765–73.

Barttfeld, P., Bekinschtein, T.A., Salles, A., Stamatakis, E.A., Adapa, R., Menon, D.K., Sigman, M., 2015. Factoring the brain signatures of anesthesia concentration and level of arousal across individuals. NeuroImage Clin. 9, 385–391.

Bassett, D.S., Bullmore, E., 2006. Small-world brain networks. Neuroscientist 12, 512–23.

Bassett, D.S., Bullmore, E.T., 2016. Small-World Brain Networks Revisited. Neurosci. 107385841666772.

Bassett, D.S., Wymbs, N.F., Porter, M. a, Mucha, P.J., Carlson, J.M., Grafton, S.T., 2011. Dynamic reconfiguration of human brain networks during learning. Proc. Natl. Acad. Sci. U. S. A. 108, 7641–6.

Bayne, T., 2018. On the axiomatic foundations of the integrated information theory of consciousness. Neurosci. Conscious. 2018, 1–8.

Bayne, T., Hohwy, J., Owen, A.M., 2016. Are There Levels of Consciousness? Trends Cogn. Sci. 20, 405–413.

Bear, M.F., Connors, B.W., Paradiso, 1996. Neuroscience: Exploring the Brain. Williams & Wilkins.

Bekinschtein, T., Cologan, V., Dahmen, B., Golombek, D., 2009. You are only coming through in waves: wakefulness variability and assessment in patients with impaired consciousness. Prog. Brain Res. 177, 171–189.

Bell, A., Sejnowski, T.J., 1995. Fast blind separation based on information theory. Proc. Intern. Symp. Nonlinear Theory Appl. 1, 43–47.

Bendat, J.S., Piersol, A.G., 1986. Random Data: Analysis and Measurement Procedures, in: Random Data: Analysis and Measurement Procedures. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp. 473–503.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B.

Berger, H., 1930. Uber das Elektrenkephalogramm des Menschen, 2nd report. J Psychol Neurol 40, 160–179.

Berger, R.J., Oswald, I., 1962. Effects of Sleep Deprivation on Behaviour, Subsequent Sleep, and Dreaming. Br. J. Psychiatry 108, 457–465.

Bigelow, H.J., 1846. Insensibility during Surgical Operations Produced by Inhalation. Bost. Med. Surg. J. 35, 309–317.

Blackman, R.B., Tukey, J.W., 1958. The Measurement of Power Spectra from the Point of View of

Communications Engineering — Part II. Bell Syst. Tech. J. 37, 485–569.

Block, N., 2009. Comparing the major theories of consciousness, in: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences. MIT Press, Cambridge, MA.

Block, N., 1996. How can we find the neural correlate of consciousness? Trends Neurosci. 19, 456–9.

Block, N., 1995. On A Confusion About a Function of Consciousness. Behav. Brain Sci. 18, 227–287.

Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. 2008, P10008.

Bodien, Y.G., Carlowicz, C.A., Chatelle, C., Giacino, J.T., 2016. Sensitivity and Specificity of the Coma Recovery Scale–Revised Total Score in Detection of Conscious Awareness. Arch. Phys. Med. Rehabil. 97, 490–492.e1.

Boly, M., Massimini, M., Garrido, M.I., Gosseries, O., Noirhomme, Q., Laureys, S., Soddu, A., 2012. Brain Connectivity in Disorders of Consciousness. Brain Connect. 2, 1–10.

Boly, M., Massimini, M., Tsuchiya, N., Postle, B.R., Koch, C., Tononi, G., 2017. Are the Neural Correlates of Consciousness in the Front or in the Back of the Cerebral Cortex? Clinical and Neuroimaging Evidence. J. Neurosci. 37, 9603–9613.

Boly, M., Sanders, R.D., Mashour, G. a, Laureys, S., 2013a. Consciousness and responsiveness: lessons from anaesthesia and the vegetative state. Curr Opin Anaesthesiol 26, 444–449.

Boly, M., Seth, A.K., Wilke, M., Ingmundson, P., Baars, B., Laureys, S., Edelman, D.B., Tsuchiya, N., 2013b. Consciousness in humans and non-human animals: recent advances and future directions. Front. Psychol. 4, 1–20.

Bonhomme, V., Boveroux, P., Brichant, J.F., Laureys, S., Boly, M., 2012. Neural correlates of consciousness during general anesthesia using functional magnetic resonance imaging (fMRI). Arch. Ital. Biol. 150, 155–163.

Bor, D., Seth, A.K., 2012. Consciousness and the prefrontal parietal network: Insights from attention, working memory, and chunking. Front. Psychol. 3, 1–14.

Boveroux, P., Vanhaudenhuyse, A., Bruno, M.-A., Noirhomme, Q., Lauwick, S., Luxen, A., Degueldre, C., Plenevaux, A., Schnakers, C., Phillips, C., Brichant, J.-F., Bonhomme, V., Maquet, P., Greicius, M.D., Laureys, S., Boly, M., 2010a. Breakdown of within- and between-network resting state functional magnetic resonance imaging connectivity during propofol-induced loss of consciousness. Anesthesiology 113, 1038–1053.

Boveroux, P., Vanhaudenhuyse, A., Phillips, C., 2010b. Breakdown of within- and between-network Resting State during Propofol-induced Loss of Consciousness. Anesthesiology 113, 1038–1053.

Brandes, U., 2001. A faster algorithm for betweenness centrality. J. Math. Sociol. 25, 163–177.

Brandner, B., Blagrove, M., McCallum, G., Bromley, L.M., 1997. Dreams, images and emotions associated with propofol anaesthesia. Anaesthesia 52, 750–5.

Bricolo, A., Turazzi, S., Faccioli, F., Odorizzi, F., Sciarretta, G., Erculiani, P., 1978. Clinical application of compressed spectral array in long-term EEG monitoring of comatose patients. Electroencephalogr. Clin. Neurophysiol. 45, 211–225.

Brigham, E.O., 1988. The Fast Fourier Transform and Its Applications. Prentice-Hall, Inc., NJ, USA.

Britz, J., 2015. Poster presented, in: Association for the Scientific Study of Consciousness (ASSC). Paris.

Britz, J., Díaz Hernàndez, L., Ro, T., Michel, C.M., 2014. EEG-microstate dependent emergence of perceptual awareness. Front. Behav. Neurosci. 8, 163.

Britz, J., Van De Ville, D., Michel, C.M., 2010. BOLD correlates of EEG topography reveal rapid resting-state network dynamics. Neuroimage 52, 1162–1170.

Brodbeck, V., Kuhn, A., von Wegner, F., Morzelewski, A., Tagliazucchi, E., Borisov, S., Michel, C.M., Laufs, H.,

2012. EEG microstates of wakefulness and NREM sleep. Neuroimage 62, 2129–2139.

Broughton, R., Hasan, J., 1995. Quantitative topographic electroencephalographic mapping during drowsiness and sleep onset. J. Clin. Neurophysiol. 12, 372–386.

Brown, E.N., Purdon, P.L., Van Dort, C.J., 2011. General Anesthesia and Altered States of Arousal: A Systems Neuroscience Analysis. Annu. Rev. Neurosci. 34, 601–628.

Bryson, B., 2003. A Short History of Nearly Everything, BMJ : British Medical Journal. Doubleday.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–98.

Busch, N.A., Dubois, J., VanRullen, R., 2009. The Phase of Ongoing EEG Oscillations Predicts Visual Perception. J. Neurosci. 29, 7869–7876.

Buzsáki, G., Anastassiou, C.A., Koch, C., 2012. The origin of extracellular fields and currents — EEG, ECoG, LFP and spikes. Nat. Rev. Neurosci. 13, 407–420.

Callahan, S., Yarmus, L., Silvestri, G., Pastis, N., 2017. Correlation Between Vital Signs and Depth of Sedation Determined by Modified Observer's Assessment of Alertness and Sedation Score. Chest 152, A886.

Cantero, J.L., Atienza, M., Salas, R.M., Gómez, C.M., 1999. Brain spatial microstates of human spontaneous alpha activity in relaxed wakefulness, drowsiness period, and REM sleep. Brain Topogr. 11, 257–263.

Casali, A.G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K.R., Casarotto, S., Bruno, M.-A., Laureys, S., Tononi, G., Massimini, M., 2013. A theoretically based index of consciousness independent of sensory processing and behavior. Sci. Transl. Med. 5, 198ra105.

Chalmers, D., 2000. What is a neural correlate of consciousness ?, in: T. Metzinger (Ed.), Neural Correlates of Consciousness: Empirical and Conceptual Issues. MIT Press.

Chalmers, D., 1995. Facing up to the Problems of Consciousness. J. Conscious. Stud. 2, 200–219.

Chalmers, D.J., 1996. The conscious mind: In search of a fundamental theory. Oxford University Press.

Chatelle, C., Chennu, S., Noirhomme, Q., Cruse, D., Owen, A.M., Laureys, S., 2012. Brain–computer interfacing in disorders of consciousness. Brain Inj. 26, 1510–22.

Chennu, S., Annen, J., Wannez, S., Thibaut, A., Chatelle, C., Cassol, H., Martens, G., Schnakers, C., Gosseries, O., Menon, D., Laureys, S., 2017. Brain networks predict metabolism, diagnosis and prognosis at the bedside in disorders of consciousness. Brain 140, 2120–2132.

Chennu, S., Bekinschtein, T.A., 2012. Arousal Modulates Auditory Attention and Awareness: Insights from Sleep, Sedation, and Disorders of Consciousness. Front. Psychol. 3, 65.

Chennu, S., Finoia, P., Kamau, E., Allanson, J., Williams, G.B., Monti, M.M., Noreika, V., Arnatkeviciute, A., Canales-Johnson, A., Olivares, F., Cabezas-Soto, D., Menon, D.K., Pickard, J.D., Owen, A.M., Bekinschtein, T. a., 2014. Spectral Signatures of Reorganised Brain Networks in Disorders of Consciousness. PLoS Comput. Biol. 10, e1003887.

Chennu, S., O'Connor, S., Adapa, R., Menon, D.K., Bekinschtein, T.A., 2016a. Brain Connectivity Dissociates Responsiveness from Drug Exposure during Propofol-Induced Transitions of Consciousness. PLOS Comput. Biol. 12, e1004669.

Chennu, S., Stamatakis, E.A., Menon, D.K., 2016b. The see-saw brain: Recovering consciousness after brain injury. Lancet Neurol. 15, 781–782.

Chiappa, K.H., Hill, R.A., 1998. Evaluation and prognostication in coma. Electroencephalogr. Clin. Neurophysiol. 106, 149–55.

Christianini, N., Shawe-Taylor, J., 2000. Support vector machines. Cambridge University Press, Cambridge, UK.

Claassen, J., Velazquez, A., Meyers, E., Witsch, J., Falo, M.C., Park, S., Agarwal, S., Michael Schmidt, J., Schiff,

N.D., Sitt, J.D., Naccache, L., Sander Connolly, E., Frey, H.P., 2016. Bedside quantitative electroencephalography improves assessment of consciousness in comatose subarachnoid hemorrhage patients. Ann. Neurol. 80, 541–553.

Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. 36, 181–204.

Clark, R.E., Squire, L.R., 1998. Classical conditioning and brain systems: the role of awareness. Science 280, 77–81.

Cohen, M.X., 2014. Effects of time lag and frequency matching on phase-based connectivity. J. Neurosci. Methods.

Collerton, D., Perry, E., McKeith, I., 2005. Why people see things that are not there: a novel Perception and Attention Deficit model for recurrent complex visual hallucinations. Behav. Brain Sci. 28, 737-57; discussion 757-94.

Comsa, I.M., Bekinschtein, T.A., Chennu, S., 2018. Transient Topographical Dynamics of the Electroencephalogram Predict Brain Connectivity and Behavioural Responsiveness During Drowsiness. Brain Topogr.

Cornblatt, B.A., Keilp, J.G., 1994. Impaired attention, genetics, and the pathophysiology of schizophrenia. Schizophr. Bull. 20, 31–46.

Crick, F., Koch, C., 2003. A framework for consciousness. Nat. Neurosci. 6, 119–26.

Crick, F., Koch, C., 1998. Consciousness and neuroscience. Cereb. Cortex 8, 97–107.

Crick, F., Koch, C., 1990. Towards a neurobiological theory of consciousness. Semin. Neurosci. 2, 263–275.

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T.A., Fernández-Espejo, D., Pickard, J.D., Laureys, S., Owen, A.M., 2011. Bedside detection of awareness in the vegetative state: A cohort study. Lancet 378, 2088–2094.

Daltrozzo, J., Wioland, N., Mutschler, V., Kotchoubey, B., 2007. Predicting coma and other low responsive patients outcome using event-related brain potentials: A meta-analysis. Clin. Neurophysiol. 118, 606–614.

Damasio, A.R., 1999. The Feeling of What Happens: Body and Emotion in the Making of Consciousness. Harcourt Brace, New York.

Darwin, C.R., 1859. On the Origin of Species. John Murray, London.

Davis, J., Goadrich, M., 2006. The Relationship Between Precision-Recall and ROC Curves. Proc. 23rd Int. Conf. Mach. Learn. -- ICML'06 233–240.

De Gennaro, L., Ferrara, M., Curcio, G., Cristiani, R., 2016. Antero-posterior EEG changes during the wakefulness and sleep transition. Clin. Neurophysiol. 112, 1901–1911.

De Gennaro, L., Vecchio, F., Ferrara, M., Curcio, G., Rossini, P.M., Babiloni, C., 2004. Changes in fronto-posterior functional coupling at sleep onset in humans. J. Sleep Res. 13, 209–217.

de Vivo, L., Bellesi, M., Marshall, W., Bushong, E.A., Ellisman, M.H., Tononi, G., Cirelli, C., 2017. Ultrastructural evidence for synaptic scaling across the wake/sleep cycle. Science (80-.). 355, 507–510.

Dehaene, S., Changeux, J.P., 2011. Experimental and Theoretical Approaches to Conscious Processing. Neuron 70, 200–227.

Dehaene, S., Charles, L., King, J.R., Marti, S., 2014. Toward a computational theory of conscious processing. Curr. Opin. Neurobiol. 25, 76–84.

Dehaene, S., Naccache, L., 2001. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. Cognition 79, 1–37.

Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

Dement, W., Kleitman, N., 1957. Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. Electroencephalogr. Clin. ….

Demertzi, A., Liew, C., Ledoux, D., Bruno, M.A., Sharpe, M., Laureys, S., Zeman, A., 2009. Dualism persists in the science of mind. Ann. N. Y. Acad. Sci. 1157, 1–9.

Dennett, D., 2007. What RoboMary Knows, in: Phenomenal Concepts and Phenomenal Knowledge. Oxford University Press, pp. 15–31.

Dennett, D.C., 1993. Consciousness explained, Artificial Intelligence. Penguin Books, London, UK.

Dennett, D.C., 1988. Quining qualia. Conscious. Mod. Sci. 1–26.

Descartes, R., 1641. Meditations on First Philosophy. SparkNotes LLC, New York.

Eastwood, J.D., Frischen, A., Fenske, M.J., Smilek, D., 2012. The Unengaged Mind. Perspect. Psychol. Sci. 7, 482–495.

Edelman, G., 2001. Consciousness: the remembered present. Ann. N. Y. Acad. Sci. 929, 111–22.

Edlow, B.L., Giacino, J.T., Wu, O., 2013a. Functional MRI and Outcome in Traumatic Coma. Curr. Neurol. Neurosci. Rep. 13, 375.

Edlow, B.L., Haynes, R.L., Takahashi, E., Klein, J.P., Cummings, P., Benner, T., Greer, D.M., Greenberg, S.M., Wu, O., Kinney, H.C., Folkerth, R.D., 2013b. Disconnection of the ascending arousal system in traumatic coma. J. Neuropathol. Exp. Neurol. 72, 505–523.

Eliade, M., 1985. A history of religious ideas: from Muhammad to the Age of Reforms. 1983 3, 360.

Eliade, M., 1982. A history of religious ideas: from Gautama Buddha to the Triumph of Chirstianity, 1978.

Eliade, M., 1981. A history of religious ideas: from the Stone Age to the Eleusinian Mysteries, 1971.

Engemann, D.A., Raimondo, F., King, J.-R., Rohaut, B., Louppe, G., Faugeras, F., Annen, J., Cassol, H., Gosseries, O., Fernandez-Slezak, D., Laureys, S., Naccache, L., Dehaene, S., Sitt, J.D., 2018. Robust EEG-based cross-site and cross-protocol classification of states of consciousness. Brain 141, 3179–3192.

Estraneo, A., Moretta, P., Loreto, V., Lanzillo, B., Cozzolino, A., Saltalamacchia, A., Lullo, F., Santoro, L., Trojano, L., 2013. Predictors of recovery of responsiveness in prolonged anoxic vegetative state. Neurology 80, 464–470.

Farooqui, A.A., Manly, T., 2017. When Attended and Conscious Perception Deactivates Fronto-Parietal Regions. Cortex 1–14.

Fell, J., Axmacher, N., 2011. The role of phase synchronization in memory processes. Nat. Rev. Neurosci. 12, 105–118.

Fellinger, R., Klimesch, W., Schnakers, C., Perrin, F., Freunberger, R., Gruber, W., Laureys, S., Schabus, M., 2011. Cognitive processes in disorders of consciousness as revealed by EEG time-frequency analyses. Clin. Neurophysiol. 122, 2177–2184.

Ferenets, R., Lipping, T., Anier, A., Jäntti, V., Melto, S., Hovilehto, S., Jantti, V., Melto, S., Hovilehto, S., 2006. Comparison of entropy and complexity measures for the assessment of depth of sedation. IEEE Trans. Biomed. Eng. 53, 1067–1077.

Ferenets, R., Vanluchene, A., Lipping, T., Heyse, B., 2007. Behavior of Entropy / Complexity Measures of the Electroencephalogram during Propofol-induced Sedation. Anesthesiology 106, 696–706.

Fernández-Espejo, D., Owen, A., 2013. Detecting awareness after severe brain injury. Nat. Rev. Neurosci. 14, 801–809.

Ferri, R., Rundo, F., Bruni, O., Terzano, M.G., Stam, C.J., 2007. Small-world network organization of functional connectivity of EEG slow-wave activity during sleep. Clin. Neurophysiol. 118, 449–456.

Fingelkurts, A. a, Fingelkurts, A. a, Kähkönen, S., 2005. Functional connectivity in the brain--is it an elusive concept? Neurosci. Biobehav. Rev. 28, 827–36.

Franks, N.P., Zecharia, A.Y., 2011. Sleep and general anesthesia. Can. J. Anesth. 58, 139–148.

Frassle, S., Sommer, J., Jansen, A., Naber, M., Einhauser, W., 2014. Binocular Rivalry: Frontal Activity Relates to Introspection and Action But Not to Perception. J. Neurosci. 34, 1738–1747.

Freeman, L.C., 1978. Centrality in social networks conceptual clarification. Soc. Networks 1, 215–239.

Friston, K., Kiebel, S., 2009. Predictive coding under the free-energy principle. Philos. Trans. R. Soc. B Biol. Sci. 364, 1211–1221.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302.

Garcia-Romeu, A., Kersgaard, B., Addy, P.H., 2016. Clinical Applications of Hallucinogens : A Review. Exp. Clin. Psychopharmacol. 24, 229–268.

Gärtner, M., Brodbeck, V., Laufs, H., Schneider, G., 2015. A stochastic model for EEG microstate sequence analysis. Neuroimage 104, 199–208.

Gazzaniga, M.S., Ivry, R.B., Mangun, G.R., 2014. Cognitive Neuroscience: The Biology of the Mind, Fourth Edi. ed. W & W Norton & Company, Ltd., London, UK.

Giacino, J.T., Fins, J.J., Laureys, S., Schiff, N.D., 2014. Disorders of consciousness after acquired brain injury: the state of the science. Nat. Rev. Neurol. 10, 99–114.

Giacino, J.T., Kalmar, K., Whyte, J., 2004. The JFK Coma Recovery Scale-Revised: Measurement characteristics and diagnostic utility. Arch. Phys. Med. Rehabil. 85, 2020–2029.

Giacino, J.T., Kezmarsky, M.A., DeLuca, J., Cicerone, K.D., 1991. Monitoring rate of recovery to predict outcome in minimally responsive patients. Arch. Phys. Med. Rehabil. 72, 897–901.

Gibbons, J.D., Chakraborti, S., 2011. Nonparametric statistical inference. Springer.

Gibson, R.M., Fernández-Espejo, D., Gonzalez-Lara, L.E., Kwan, B.Y., Lee, D.H., Owen, A.M., Cruse, D., 2014. Multiple tasks and neuroimaging modalities increase the likelihood of detecting covert awareness in patients with disorders of consciousness. Front. Hum. Neurosci. 8, 950.

Gosseries, O., Schnakers, C., Ledoux, D., Vanhaudenhuyse, A., Bruno, M.-A., Demertzi, A., Noirhomme, Q., Lehembre, R., Damas, P., Goldman, S., Peeters, E., Moonen, G., Laureys, S., 2011. Automated EEG entropy measurements in coma, vegetative state/unresponsive wakefulness syndrome and minimally conscious state. Funct. Neurol. 26, 25–30.

Goupil, L., Bekinschtein, T. a., 2012. Cognitive processing during the transition to sleep. Arch. Ital. Biol. 150, 140–154.

Greenhouse, S.W., Geisser, S., 1959. On methods in the analysis of profile data. Psychometrika 24, 95–112.

Grieder, M., Koenig, T., Kinoshita, T., Utsunomiya, K., Wahlund, L.O., Dierks, T., Nishida, K., 2016. Discovering EEG resting state alterations of semantic dementia. Clin. Neurophysiol. 127, 2175–2181.

Gross, D.W., Gotman, J., 1999. Correlation of high-frequency oscillations with the sleep–wake cycle and cognitive activity in humans. Neuroscience 94, 1005–1018.

Gugino, L.D., Chabot, R.J., Prichep, L.S., John, E.R., Formanek, V., Aglio, L.S., 2001. Quantitative EEG changes associated with loss and return of consciousness in healthy adult volunteers anaesthetized with propofol or sevoflurane. Br. J. Anaesth. 87, 421–428.

Guimerà, R., Amaral, L.A.N., 2005. Cartography of complex networks: modules and universal roles. J. Stat. Mech. Theory Exp. 2005, P02001.

Guimera, R., Sales-Pardo, M., Amaral, L., 2004. Modularity from fluctuations in random graphs and complex networks. Phys. Rev. E.

Haas, L.F., 2003. Hans Berger (1873-1941), Richard Caton (1842-1926), and electroencephalography. J. Neurol. Neurosurg. Psychiatry 74, 9.

Hanslmayr, S., Gross, J., Klimesch, W., Shapiro, K.L., 2011. The role of alpha oscillations in temporal attention. Brain Res. Rev. 67, 331–343.

Harrison, A.H., Connolly, J.F., 2013. Finding a way in: A review and practical evaluation of fMRI and EEG for detection and assessment in disorders of consciousness. Neurosci. Biobehav. Rev. 37, 1403–1419.

Harrison, Y., Horne, J.A., 2000. The impact of sleep deprivation on decision making: A review. J. Exp. Psychol. Appl. 6, 236–249.

Hashemi, M., Hutt, A., Hight, D., Sleigh, J., 2017. Anesthetic action on the transmission delay between cortex and thalamus explains the beta-buzz observed under propofol anesthesia. PLoS One 12, 1–29.

Hassan, M., Dufor, O., Merlet, I., Berrou, C., Wendling, F., 2014. EEG source connectivity analysis: From dense array recordings to brain networks. PLoS One 9.

Hawking, S.W., 2001. The universe in a nutshell. Bandtam Books, New York.

He, B.J., Snyder, A.Z., Vincent, J.L., Epstein, A., Shulman, G.L., Corbetta, M., 2007. Breakdown of Functional Connectivity in Frontoparietal Networks Underlies Behavioral Deficits in Spatial Neglect. Neuron 53, 905–918.

Heine, L., Soddu, A., Gómez, F., Vanhaudenhuyse, A., Tshibanda, L., Thonnard, M., Charland-Verville, V., Kirsch, M., Laureys, S., Demertzi, A., 2012. Resting state networks and consciousness Alterations of multiple resting state network connectivity in physiological, pharmacological, and pathological consciousness states. Front. Psychol. 3, 1–12.

Herrmann, M., 2001. Release of biochemical markers of damage to neuronal and glial brain tissue is associated with short and long term neuropsychological outcome after traumatic brain injury. J. Neurol. Neurosurg. Psychiatry 70, 95–100.

Hippocrates, 400BC. On the Sacred Disease.

Hobson, J.A., 1995. Sleep. Scientific American Library.

Hobson, J.A., Pace-Schott, E.F., 2002. The cognitive neuroscience of sleep: neuronal systems, consciousness and learning. Nat. Rev. Neurosci. 3, 679–93.

Hogg, R. V., Ledolter, J., 1987. Engineering Statistics. MacMillan, New York.

Hohwy, J., 2009. The neural correlates of consciousness: New experimental approaches needed? Conscious. Cogn. 18, 428–438.

Holm, S., 1979. A Simple Sequentially Rejective Multiple Test Procedure. Scand. J. Stat. 6, 65–70.

Hori, T., Hayashi, M., Morikawa, T., 1994. Topographical EEG changes and the hypnagogic experience., in: Sleep Onset: Normal and Abnormal Processes. American Psychological Association, Washington, pp. 237–253.

Horne, J., Reyner, L., 1999. Vehicle accidents related to sleep: a review. Occup. Environ. Med. 56, 289–294.

Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. A Math. Phys. Eng. Sci.

Hudetz, A.G., Liu, X., Pillay, S., 2015. Dynamic Repertoire of Intrinsic Brain States Is Reduced in Propofol-Induced Unconsciousness. Brain Connect. 5, 10–22.

Hudetz, A.G., Liu, X., Pillay, S., Boly, M., Tononi, G., 2016. Propofol anesthesia reduces Lempel-Ziv complexity of spontaneous brain activity in rats. Neurosci. Lett. 628, 132–135.

Humphries, M.D., Gurney, K., 2008. Network “small-world-ness”: A quantitative method for determining canonical network equivalence. PLoS One 3.

Iber, C., Ancoli-Israel, S., Chesson Jr., A.L., Quan, S.F., 2007. The AASM Manual for the Scoring of Sleep and Associated Events: Rules Terminology and Technical Specifications 1st ed.

Jackson, F., 1982. Epiphenomenal Qualia. Philos. Q. 32, 127.

James, W., 1904. Does "Consciousness" Exist? J. Philos. Psychol. Sci. Methods 1, 477–491.

John Locke, 1689. An Essay Concerning Human Understanding.

Jones, C., 1979. Glasgow Coma Scale. Am. J. Nurs. 79, 1551.

Kajimura, N., Uchiyama, M., Takayama, Y., Uchida, S., Uema, T., Kato, M., Sekimoto, M., Watanabe, T., Nakajima, T., Horikoshi, S., Ogawa, K., Nishikawa, M., Hiroki, M., Kudo, Y., Matsuda, H., Okawa, M., Takahashi, K., 1999. Activity of midbrain reticular formation and neocortex during the progression of human non-rapid eye movement sleep. J. Neurosci. 19, 10065–73.

Kalmar, K., Giacino, J.T., 2007. The JFK Coma Recovery Scale--Revised. Neuropsychol. Rehabil. 15, 454–460.

Kane, N.M., Curry, S.H., Rowlands, C. a, Manara, a R., Lewis, T., Moss, T., Cummins, B.H., Butler, S.R., 1996. Event-related potentials--neurophysiological tools for predicting emergence and early outcome from traumatic coma. Intensive Care Med. 22, 39–46.

Kane, N.M., Moss, T.H., Curry, S.H., Butler, S.R., 1998. Quantitative electroencephalographic evaluation of non-fatal and fatal traumatic coma. Electroencephalogr. Clin. Neurophysiol. 106, 244–250.

Kaplan, P.W., 2004. The EEG in metabolic encephalopathy and coma. J. Clin. Neurophysiol. 21, 307–18.

Karnaze, D.S., Marshall, L.F., Bickford, R.G., 1982. EEG monitoring of clinical coma: the compressed spectral array. Neurology 32, 289–292.

Katayama, H., Gianotti, L.R.R., Isotani, T., Faber, P.L., Sasada, K., Kinoshita, T., Lehmann, D., 2007. Classes of multichannel EEG microstates in light and deep hypnotic conditions. Brain Topogr. 20, 7–14.

Khanna, A., Pascual-Leone, A., Farzan, F., 2014. Reliability of Resting-State Microstate Features in Electroencephalography. PLoS One 9, e114163.

Khanna, A., Pascual-Leone, A., Michel, C.M., Farzan, F., 2015. Microstates in resting-state EEG: current status and future directions. Neurosci. Biobehav. Rev. 49, 105–13.

Kikuchi, M., Koenig, T., Munesue, T., Hanaoka, A., Strik, W., Dierks, T., Koshino, Y., Minabe, Y., 2011. EEG microstate analysis in Drug-Naive patients with panic disorder. PLoS One 6, 2–7.

Kikuchi, M., Koenig, T., Wada, Y., Higashima, M., Koshino, Y., Strik, W., Dierks, T., 2007. Native EEG and treatment effects in neuroleptic-naïve schizophrenic patients: Time and frequency domain approaches. Schizophr. Res. 97, 163–172.

Kim, H., Hudetz, A.G., Lee, J., Mashour, G.A., Lee, U., 2018. Estimating the Integrated Information Measure Phi from High-Density Electroencephalography during States of Consciousness in Humans. Front. Hum. Neurosci. 12, 1–13.

Kindler, J., Hubl, D., Strik, W.K., Dierks, T., Koenig, T., 2011. Resting-state EEG in schizophrenia: Auditory verbal hallucinations are related to shortening of specific microstates. Clin. Neurophysiol. 122, 1179–1182.

King, J.R., Sitt, J.D., Faugeras, F., Rohaut, B., El Karoui, I., Cohen, L., Naccache, L., Dehaene, S., 2013. Information sharing in the brain indexes consciousness in noncommunicative patients. Curr. Biol. 23, 1914–1919.

Kirschstein, T., Köhling, R., 2009. What is the Source of the EEG? Clin. EEG Neurosci. 40, 146–149.

Klimesch, W., 2012. Alpha-band oscillations, attention, and controlled access to stored information. Trends Cogn. Sci. 16, 606–617.

Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res. Rev. 29, 169–195.

Knutson, K.L., Spiegel, K., Penev, P., Van Cauter, E., 2007. The metabolic consequences of sleep deprivation. Sleep Med. Rev. 11, 163–178.

Koch, C., Massimini, M., Boly, M., Tononi, G., 2016. Neural correlates of consciousness: progress and problems. Nat. Rev. Neurosci. 17, 307–321.

Koch, C., Tsuchiya, N., 2007. Attention and consciousness: two distinct brain processes. Trends Cogn. Sci. 11, 16–22.

Koenig, T., Brandeis, D., 2016. Inappropriate assumptions about EEG state changes and their impact on the quantification of EEG state dynamics. Neuroimage 125, 1104–1106.

Koenig, T., Lehmann, D., Merlo, M.C.G., Kochi, K., Hell, D., Koukkou, M., 1999. A deviant EEG brain microstate in acute, neuroleptic-naive schizophrenics at rest. Eur. Arch. Psychiatry Clin. Neurosci. 249, 205–211.

Koenig, T., Prichep, L., Lehmann, D., Sosa, P.V., Braeker, E., Kleinlogel, H., Isenhart, R., John, E.R., 2002. Millisecond by millisecond, year by year: normative EEG microstates and developmental stages. Neuroimage 16, 41–48.

Koenig, T., Studer, D., Hubl, D., Melie, L., Strik, W.K., 2005. Brain connectivity at different time-scales measured with EEG. Philos. Trans. R. Soc. B Biol. Sci. 360, 1015–1024.

Kotchoubey, B., Lang, S., Baales, R., Herb, E., Maurer, P., Mezger, G., Schmalohr, D., Bostanov, V., Birbaumer, N., 2001. Brain potentials in human patients with extremely severe diffuse brain damage. Neurosci. Lett. 301, 37–40.

Kouider, S., Andrillon, T., Barbosa, L.S., Goupil, L., Bekinschtein, T.A., 2014. Inducing Task-Relevant Responses to Speech in the Sleeping Brain. Curr. Biol. 24, 2208–2214.

Ku, S.W., Lee, U., Noh, G.J., Jun, I.G., Mashour, G. a., 2011. Preferential inhibition of frontal-to-parietal feedback connectivity is a neurophysiologic correlate of general anesthesia in surgical patients. PLoS One 6, 1–9.

Kuhn, A., Brodbeck, V., Tagliazucchi, E., Morzelewski, A., von Wegner, F., Laufs, H., 2015. Narcoleptic Patients Show Fragmented EEG-Microstructure During Early NREM Sleep. Brain Topogr. 28, 619–635.

Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J., 1999. Measuring phase synchrony in brain signals. Hum. Brain Mapp. 8, 194–208.

Lamme, V.A.F., 2003. Why visual attention and awareness are different. Trends Cogn. Sci. 7, 12–18.

Larkum, M.E., Nevian, T., 2008. Synaptic clustering by dendritic signalling mechanisms. Curr. Opin. Neurobiol. 18, 321–331.

Larson-Prior, L.J., Power, J.D., Vincent, J.L., Nolan, T.S., Coalson, R.S., Zempel, J., Snyder, A.Z., Schlaggar, B.L., Raichle, M.E., Petersen, S.E., 2011. Modulation of the brain's functional network architecture in the transition from wake to sleep. Prog. Brain Res. 193, 277–294.

Larson-Prior, L.J., Zempel, J.M., Nolan, T.S., Prior, F.W., Snyder, A.Z., Raichle, M.E., 2009. Cortical network functional connectivity in the descent to sleep. Proc. Natl. Acad. Sci. 106, 4489–4494.

Larson, E.J., Witham, L., 1997. Scientists are still keeping the faith. Nature 386, 435–436.

Laureys, S., 2005. The neural correlate of (un)awareness: lessons from the vegetative state. Trends Cogn. Sci. 9, 556–559.

Laureys, S., Perrin, F., Schnakers, C., Boly, M., Majerus, S., 2005. Residual cognitive function in comatose, vegetative and minimally conscious states. Curr. Opin. Neurol. 18, 726–733.

Laureys, S., Schiff, N.D., 2012. Coma and consciousness: Paradigms (re)framed by neuroimaging. Neuroimage 61, 478–491.

Laureys, S., Tononi, G., 2011. The neurology of consciousness: cognitive neuroscience and neuropathology.

Lechinger, J., Bothe, K., Pichler, G., Michitsch, G., Donis, J., Klimesch, W., Schabus, M., 2013. CRS-R score in

disorders of consciousness is strongly related to spectral EEG at rest. J. Neurol. 260, 2348–2356.

Lee, H., Mashour, G. a, Noh, G.-J., Kim, S., Lee, U., 2013. Reconfiguration of network hub structure after propofol-induced unconsciousness. Anesthesiology 119, 1347–59.

Lee, H., Noh, G.-J.J., Joo, P., Choi, B.-M.M., Silverstein, B.H., Kim, M., Wang, J., Jung, W.-S.S., Kim, S., 2017. Diversity of functional connectivity patterns is reduced in propofol-induced unconsciousness. Hum. Brain Mapp. 38, 4980–4995.

Lee, U., Kim, S., Noh, G.J., Choi, B.M., Hwang, E., Mashour, G. a., 2009a. The directionality and functional organization of frontoparietal connectivity during consciousness and anesthesia in humans. Conscious. Cogn. 18, 1069–1078.

Lee, U., Ku, S., Noh, G., Baek, S., Choi, B., Mashour, G. a, 2013. Disruption of Frontal-Parietal Communication by Ketamine, Propofol, and Sevoflurane. Anesthesiology 118, 1264–1275.

Lee, U., Mashour, G. a., Kim, S., Noh, G.J., Choi, B.M., 2009b. Propofol induction reduces the capacity for neural information integration: Implications for the mechanism of consciousness and general anesthesia. Conscious. Cogn. 18, 56–64.

Lehembre, R., Bruno, M.A., Vanhaudenhuyse, A., Chatelle, C., Cologan, V., Leclercq, Y., Soddu, A., Macq, B., Laureys, S., Noirhomme, Q., 2012a. Resting-state EEG study of comatose patients: A connectivity and frequency analysis to find differences between vegetative and minimally conscious states. Funct. Neurol. 27, 41–47.

Lehembre, R., Gosseries, O., Lugo, Z., Jedidi, Z., Chatelle, C., Sadzot, B., Laureys, S., Noirhomme, Q., 2012b. Electrophysiological investigations of brain function in coma, vegetative and minimally conscious patients. Arch. Ital. Biol. 150, 122–139.

Lehmann, D., 1990. Brain electric microstates and cognition: the atoms of thought, in: Machinery of the Mind. Springer, pp. 209–224.

Lehmann, D., 1971. Multichannel topography of human alpha EEG fields. Electroencephalogr. Clin. Neurophysiol. 31, 439–449.

Lehmann, D., Faber, P.L., Galderisi, S., Herrmann, W.M., Kinoshita, T., Koukkou, M., Mucci, A., Pascual-Marqui, R.D., Saito, N., Wackermann, J., Winterer, G., Koenig, T., 2005. EEG microstate duration and syntax in acute, medication-naïve, first-episode schizophrenia: A multi-center study. Psychiatry Res. - Neuroimaging 138, 141–156.

Lehmann, D., Ozaki, H., Pal, I., 1987. EEG alpha map series: brain micro-states by space-oriented adaptive segmentation. Electroencephalogr. Clin. Neurophysiol. 67, 271–288.

Lehmann, D., Pascual-Marqui, R.D., Strik, W.K., Koenig, T., 2010. Core networks for visual-concrete and abstract thought content: A brain electric microstate analysis. Neuroimage 49, 1073–1079.

Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr. Clin. Neurophysiol. 48, 609–621.

Lempel, A., Ziv, J., 1976. On the Complexity of Finite Sequences. IEEE Trans. Inf. Theory 22, 75–81.

Leslie, K., Sleigh, J., Paech, M.J., Voss, L., Lim, C.W., Sleigh, C., 2009. Dreaming and electroencephalographic changes during anesthesia maintained with propofol or desflurane. Anesthesiology 111, 547–555.

Levine, J., 1983. Materialism and qualia: the explanatory gap. Pacific Philos. Q.

Lewis, D., 1999. What experience teaches, in: Papers in Metaphysics and Epistemology. Cambridge University Press, Cambridge, pp. 262–290.

Lewis, L.D., Weiner, V.S., Mukamel, E.A., Donoghue, J.A., Eskandar, E.N., Madsen, J.R., Anderson, W.S., Hochberg, L.R., Cash, S.S., Brown, E.N., Purdon, P.L., 2012. Rapid fragmentation of neuronal networks at the onset of propofol-induced unconsciousness. Proc. Natl. Acad. Sci. 109, E3377–E3386.

Loomis, A.L., Harvey, E.N., Hobart, G., 1935. Potential rhythms of the cerebral cortex during sleep. Science 81, 597–598.

Lynall, M.-E., Bassett, D.S., Kerwin, R., McKenna, P.J., Kitzbichler, M., Muller, U., Bullmore, E., 2010. Functional connectivity and brain networks in schizophrenia. J. Neurosci. 30, 9477–87.

Lyratzopoulos (Liratsopulos), G., Blain, K.M., 2003. Inhalation sedation with nitrous oxide as an alternative to dental general anaesthesia for children. J. Public Health (Bangkok). 25, 303–312.

MacDonald, A.A., Naci, L., MacDonald, P.A., Owen, A.M., 2015. Anesthesia and neuroimaging: Investigating the neural correlates of unconsciousness. Trends Cogn. Sci. 19, 100–107.

Marchant, N., Sanders, R., Sleigh, J., Vanhaudenhuyse, A., Bruno, M.-A., Brichant, J.F., Laureys, S., Bonhomme, V., 2014. How electroencephalography serves the anesthesiologist. Clin. EEG Neurosci. 45, 22–32.

Marple, L., 1999. Computing the discrete-time "analytic" signal via FFT. IEEE Trans. Signal Process. 47, 2600–2603.

Mashour, G.A., 2004. Consciousness unbound: Toward a paradigm of general anesthesia. Anesthesiology 100, 428–433.

Mashour, G.A., Avidan, M.S., 2017. Black swans: challenging the relationship of anaesthetic-induced unconsciousness and electroencephalographic oscillations in the frontal cortex. BJA Br. J. Anaesth. 12, 1–29.

Massey, F.J., 1951. The Kolmogorov-Smirnov Test for Goodness of Fit. J. Am. Stat. Assoc. 46, 68–78.

Massimini, M., Ferrarelli, F., Esser, S.K., Riedner, B.A., Huber, R., Murphy, M., Peterson, M.J., Tononi, G., 2007. Triggering sleep slow waves by transcranial magnetic stimulation. Proc. Natl. Acad. Sci. 104, 8496–8501.

Massimini, M., Ferrarelli, F., Huber, R., Esser, S.K., Singh, H., Tononi, G., 2005. Breakdown of cortical effective connectivity during sleep. Science 309, 2228–2232.

Mathewson, K.E., Gratton, G., Fabiani, M., Beck, D.M., Ro, T., 2009. To See or Not to See: Prestimulus Phase Predicts Visual Awareness. J. Neurosci. 29, 2725–2732.

Mauchly, J.W., 1940. Significance Test for Sphericity of a Normal n-Variate Distribution. Ann. Math. Stat. 11, 204–209.

Maxwell, J.A., 2004. Using Qualitative Methods for Causal Explanation. Field methods 16, 243–264.

McGinn, C., 1989. Can we solve the mind-body problem? Mind 98, 349–366.

Mediano, P.A.M., Seth, A.K., Barrett, A.B., 2018. Measuring Integrated Information: Comparison of Candidate Measures in Theory and Simulation 1–30.

Menon, D.K., 1999. Cerebral protection in severe brain injury: physiological determinants of outcome and their optimisation. Br. Med. Bull. 55, 226–258.

Metzinger, T., 2000. Neural correlates of consciousness: Empirical and conceptual questions. MIT press.

Michel, C.M., Koenig, T., Brandeis, D., 2009. Electrical neuroimaging in the time domain, in: Michel, C.M., Koenig, T., Brandeis, D., Gianotti, L.R.R., Wackermann, J. (Eds.), Electrical Neuroimaging. Cambridge University Press, Cambridge, pp. 111–144.

Milgram, S., 1967. The small world problem. Psychol. Today 1, 61–67.

Milz, P., Faber, P.L., Lehmann, D., Koenig, T., Kochi, K., Pascual-Marqui, R.D., 2015. The functional significance of EEG microstates—Associations with modalities of thinking. Neuroimage 125, 643–656.

Milz, P., Pascual-Marqui, R.D., Achermann, P., Kochi, K., Faber, P.L., 2017. The EEG microstate topography is predominantly determined by intracortical sources in the alpha band. Neuroimage 162, 353–361.

Monti, M., Vanhaudenhuyse, A., Coleman, M., Boly, M., Pickard, J., Tshibanda, L., Owen, A., Laureys, S., 2010.

Willful modulation of brain activity in disorders of consciousness. Engl. J.

Monti, M.M., Lutkenhoff, E.S., Rubinov, M., Boveroux, P., Vanhaudenhuyse, A., Gosseries, O., Bruno, M.A., Noirhomme, Q., Boly, M., Laureys, S., 2013. Dynamic Change of Global and Local Information Processing in Propofol-Induced Loss and Recovery of Consciousness. PLoS Comput. Biol. 9.

Morikawa, T., Hayashi, M., Hori, T., 1997. Auto power and coherence analysis of delta-theta band EEG during the waking-sleeping transition period. Electroencephalogr. Clin. Neurophysiol. 103, 633–641.

Murphy, M., Bruno, M.-A., Riedner, B.A., Boveroux, P., Noirhomme, Q., Landsness, E.C., Brichant, J.-F., Phillips, C., Massimini, M., Laureys, S., Tononi, G., Boly, M., 2011. Propofol anesthesia and sleep: a high-density EEG study. Sleep 34, 283–291A.

Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: A step-by-step tutorial review. Brain Topogr. 20, 249–264.

Naccache, L., Puybasset, L., Gaillard, R., Serve, E., Willer, J.C., 2005. Auditory mismatch negativity is a good predictor of awakening in comatose patients: A fast and reliable procedure [1]. Clin. Neurophysiol. 116, 988–989.

Naci, L., Cusack, R., Anello, M., Owen, A.M., 2014. A common neural code for similar conscious experiences in different individuals. Proc. Natl. Acad. Sci. U. S. A. 111, 14277–14282.

Nagel, T., 1974. What Is It Like to Be a Bat? Philos. Rev. 83, 435.

Naghavi, H.R., Nyberg, L., 2005. Common fronto-parietal activity in attention, memory, and consciousness: Shared demands on integration? Conscious. Cogn. 14, 390–425.

Newman, M., Girvan, M., 2004. Finding and evaluating community structure in networks. Phys. Rev. E 69, 1–16.

Niedermeyer, E., 2005a. The Normal EEG of the Waking Adult, in: Niedermeyer, E., Lopes Da Silva, F. (Eds.), Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams & Wilkins, Philadelphia.

Niedermeyer, E., 2005b. Sleep and EEG, in: Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams & Wilkins, Philadelphia.

Niedermeyer, E., Lopes Da Silva, F., 2005. Electroencephalography: Basic Principles, Clinical Applications, Related Fields, Fifth. ed. Lippincott Williams & Wilkins, Philadelphia.

Nishida, K., Morishima, Y., Yoshimura, M., Isotani, T., Irisawa, S., Jann, K., Dierks, T., Strik, W., Kinoshita, T., Koenig, T., 2013. EEG microstates associated with salience and frontoparietal networks in frontotemporal dementia, schizophrenia and Alzheimer's disease. Clin. Neurophysiol. 124, 1106–1114.

Noreika, V., Canales-Johnson, A., Koh, J., Taylor, M., Massey, I., Bekinschtein, T. a., 2015. Intrusions of a drowsy mind: neural markers of phenomenological unpredictability. Front. Psychol. 6, 1–10.

Nunez, P.L., Srinivasan, R., 2006. Electric fields of the brain: the neurophysics of EEG. Oxford university press.

Nunez, P.L., Wingeier, B.M., Silberstein, R.B., 2001. Spatial-temporal structures of human alpha rhythms: Theory, microcurrent sources, multiscale measurements, and global binding of local networks. Hum. Brain Mapp. 13, 125–164.

Odegaard, B., Knight, R.T., Lau, H., 2017. Should a Few Null Findings Falsify Prefrontal Theories of Conscious Perception? J. Neurosci. 37, 9593–9602.

Ogilvie, R.D., 2001. The process of falling asleep. Sleep Med. Rev. 5, 247–270.

Ogilvie, R.D., Wilkinson, R.T., 1984. The detection of sleep onset: behavioral and physiological convergence. Psychophysiology 21, 510–520.

Oizumi, M., Albantakis, L., Tononi, G., 2014. From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. PLoS Comput. Biol. 10.

Oizumi, M., Amari, S.I., Yanagawa, T., Fujii, N., Tsuchiya, N., 2016. Measuring Integrated Information from the Decoding Perspective. PLoS Comput. Biol. 12, 1–18.

Onnela, J.P., Saramäki, J., Kertész, J., Kaski, K., 2005. Intensity and coherence of motifs in weighted complex networks. Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys. 71.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011.

Oro, J., Miller, S.L., Lazcano, A., 1990. The Origin and Early Evolution of Life on Earth. Annu. Rev. Earth Planet. Sci. 18, 317–356.

Overgaard, M., 2009. How can we know if patients in coma, vegetative state or minimally conscious state are conscious? Prog. Brain Res. 177, 11–19.

Overgaard, M., Overgaard, R., 2011. Measurements of consciousness in the vegetative state. Lancet 378, 2052–4.

Overgaard, M., Overgaard, R., 2010. Neural correlates of contents and levels of consciousness. Front. Psychol. 1, 164.

Owen, A.M., Coleman, M.R., 2008. Detecting awareness in the vegetative state, in: Annals of the New York Academy of Sciences. pp. 130–138.

Owen, A.M., Coleman, M.R., Boly, M., Davis, M.H., Laureys, S., Pickard, J.D., 2006. Detecting Awareness in the Vegetative State. Science (80-.). 313, 1402–1402.

Palva, S., Palva, J.M., 2011. Functional Roles of Alpha-Band Phase Synchronization in Local and Large-Scale Cortical Networks. Front. Psychol. 2, 1–15.

Pandit, J.J., Andrade, J., Bogod, D.G., Hitchman, J.M., Jonker, W.R., Lucas, N., Mackay, J.H., Nimmo, A.F., O'Connor, K., O'Sullivan, E.P., Paul, R.G., Palmer, J.H.M.G., Plaat, F., Radcliffe, J.J., Sury, M.R.J., Torevell, H.E., Wang, M., Hainsworth, J., Cook, T.M., Armstrong, J., Bird, J., Eddy, A., Harrop-Griffiths, W., Love, N., Mahajan, R., Mallick, A., Barker, I., Kirkpatrick, A., Molodynski, J., Poonnusamy, K., Moonesinghe, R., Weatherill, D., Smith, D., Grocott, M., Humphrey, M., Casserly, M., Drake, S., Rangasami, J., 2014. 5th National Audit Project (NAP5) on accidental awareness during general anaesthesia: Summary of main findings and risk factors. Br. J. Anaesth. 113, 549–559.

Pascual-Marqui, R.D., Lehmann, D., Faber, P., Milz, P., Kochi, K., Yoshimura, M., Nishida, K., Isotani, T., Kinoshita, T., 2014. The resting microstate networks (RMN): cortical distributions, dynamics, and frequency specific information flow. arXiv Prepr. arXiv1411.1949 1–14.

Pasqual-Marqui, R.D., Michel, C.M., Lehmann, D., 1995. Segmentation of brain electrical activity into microstates: model estimation and validation. IEEE Trans. Biomed. Eng. 42, 658–665.

Pfurtscheller, G., Stancák, A., Neuper, C., 1996. Event-related synchronization (ERS) in the alpha band — an electrophysiological correlate of cortical idling: A review. Int. J. Psychophysiol. 24, 39–46.

Pilcher, J.J., Huffcutt, A.I., 1996. Effects of sleep deprivation on performance: A meta-analysis. Sleep 19, 318–326.

Pinto, Y., de Haan, E.H.F., Lamme, V.A.F., 2017. The Split-Brain Phenomenon Revisited: A Single Conscious Agent with Split Perception. Trends Cogn. Sci. 21, 835–851.

Platt, J.C., 1999. Probabilistic Outputs for Support Vector Machines and Comparisons to Regular Likelihood Methods. Adv. Large Margin Classif.

Prerau, M.J., Hartnack, K.E., Obregon-Henao, G., Sampson, A., Merlino, M., Gannon, K., Bianchi, M.T., Ellenbogen, J.M., Purdon, P.L., 2014. Tracking the Sleep Onset Process: An Empirical Model of Behavioral and Physiological Dynamics. PLoS Comput. Biol. 10, e1003866.

Purdon, P.L., Pierce, E.T., Mukamel, E. a, Prerau, M.J., Walsh, J.L., Wong, K.F.K., Salazar-Gomez, A.F., Harrell, P.G., Sampson, A.L., Cimenser, A., Ching, S., Kopell, N.J., Tavares-Stoeckel, C., Habeeb, K., Merhar, R.,

Brown, E.N., 2013. Electroencephalogram signatures of loss and recovery of consciousness from propofol. Proc. Natl. Acad. Sci. U. S. A. 110, E1142-51.

Rampersad, S.E., Mulroy, M.F., 2005. A case of awareness despite an "adequate depth of anesthesia" as indicated by a Bispectral Index?? monitor. Anesth. Analg. 100, 1363–1364.

Rees, G., Kreiman, G., Koch, C., 2002. Neural correlates of consciousness in humans. Nat. Rev. Neurosci. 3, 261–270.

Reijneveld, J.C., Ponten, S.C., Berendse, H.W., Stam, C.J., 2007. The application of graph theoretical analysis to complex networks in the brain. Clin. Neurophysiol. 118, 2317–2331.

Ronzhina, M., Janoušek, O., Kolářová, J., Nováková, M., Honzík, P., Provazník, I., 2012. Sleep scoring using artificial neural networks. Sleep Med. Rev. 16, 251–263.

Rosenthal, D.M., 2009. Concepts and Definitions of Consicousness, in: William P. Banks (Ed.), Encyclopedia of Consciousness. Elsevier, Amsterdam.

Rubinov, M., Knock, S.A., Stam, C.J., Micheloyannis, S., Harris, A.W.F., Williams, L.M., Breakspear, M., 2009. Small-world properties of nonlinear brain activity in schizophrenia. Hum. Brain Mapp. 30, 403–16.

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52, 1059–69.

Rumpl, E., Prugger, M., Bauer, G., Gerstenbrand, F., Hackl, J.M., Pallua, A., 1983. Incidence and prognostic value of spindles in post-traumatic coma. Electroencephalogr. Clin. Neurophysiol. 56, 420–429.

Russell, I.F., 2013. The ability of bispectral index to detect intra-operative wakefulness during isoflurane/air anaesthesia, compared with the isolated forearm technique. Anaesthesia 68, 1010–1020.

Sakkalis, V., 2011. Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. Comput. Biol. Med. 41, 1110–1117.

Samaha, J., 2015. How best to study the function of consciousness? Front. Psychol. 6, 1–3.

Samu, D., Seth, A.K., Nowotny, T., 2014. Influence of Wiring Cost on the Large-Scale Architecture of Human Cortical Connectivity. PLoS Comput. Biol. 10, e1003557.

Sanders, R.D., Tononi, G., Laureys, S., Sleigh, J.W., 2012. Unresponsiveness ≠ unconsciousness. Anesthesiology 116, 946–59.

Sarasso, S., Boly, M., Napolitani, M., Gosseries, O., Charland-Verville, V., Casarotto, S., Rosanova, M., Casali, A.G., Brichant, J.F., Boveroux, P., Rex, S., Tononi, G., Laureys, S., Massimini, M., 2015. Consciousness and complexity during unresponsiveness induced by propofol, xenon, and ketamine. Curr. Biol. 25, 3099–3105.

Sauseng, P., Klimesch, W., 2008. What does phase information of oscillatory brain activity tell us about cognitive processes? Neurosci. Biobehav. Rev. 32, 1001–1013.

Scammell, T.E., Arrigoni, E., Lipton, J.O., 2017. Neural Circuitry of Wakefulness and Sleep. Neuron 93, 747–765.

Schartner, M.M., Carhart-Harris, R.L., Barrett, A.B., Seth, A.K., Muthukumaraswamy, S.D., 2017a. Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, LSD and psilocybin. Sci. Rep. 7, 46421.

Schartner, M.M., Pigorini, A., Gibbs, S.A., Arnulfo, G., Sarasso, S., Barnett, L., Nobili, L., Massimini, M., Seth, A.K., Barrett, A.B., 2017b. Global and local complexity of intracranial EEG decreases during NREM sleep. Neurosci. Conscious. niw022.

Schartner, M.M., Seth, A., Noirhomme, Q., Boly, M., Bruno, M.-A., Laureys, S., Barrett, A., 2015. Complexity of Multi-Dimensional Spontaneous EEG Decreases during Propofol Induced General Anaesthesia. PLoS One 10, e0133532.

Schiff, N.D., Nauvel, T., Victor, J.D., 2014. Large-scale brain dynamics in disorders of consciousness. Curr. Opin. Neurobiol. 25, 7–14.

Schlegel, F., Lehmann, D., Faber, P.L., Milz, P., Gianotti, L.R.R., 2012. EEG Microstates During Resting Represent Personality Differences. Brain Topogr. 25, 20–26.

Schnakers, C., Majerus, S., Giacino, J., Vanhaudenhuyse, A., Bruno, M.-A., Boly, M., Moonen, G., Damas, P., Lambermont, B., Lamy, M., Damas, F., Ventura, M., Laureys, S., 2008. A French validation study of the Coma Recovery Scale-Revised (CRS-R). Brain Inj. 22, 786–792.

Schnakers, C., Vanhaudenhuyse, A., Giacino, J., Ventura, M., Boly, M., Majerus, S., Moonen, G., Laureys, S., 2009. Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. BMC Neurol. 9, 35.

Schoffelen, J.M., Gross, J., 2009. Source connectivity analysis with MEG and EEG. Hum. Brain Mapp. 30, 1857–1865.

Schreiber, T., 2000. Measuring information transfer. Phys. Rev. Lett. 85, 461–464.

Schwartz, R.S., Brown, E.N., Lydic, R., Schiff, N.D., 2010. General Anesthesia, Sleep, and Coma. N. Engl. J. Med. 363, 2638–2650.

Searle, J.R., 1997. The Mystery Of Consciousness. New York Review of Books.

Searle, J.R., 1993. The problem of consciousness. Ciba Found. Symp. 174, 61-9; discussion 70-80.

Sebel, P.S., Bowdle, T.A., Ghoneim, M.M., Rampil, I.J., Padilla, R.E., Gan, T.J., Domino, K.B., 2004. The incidence of awareness during anesthesia: A multicenter United States study. Anesth. Analg. 99, 833–839.

Seifert, H. a, Blouin, R.T., Conard, P.F., Gross, J.B., 1993. Sedative doses of propofol increase beta activity of the processed electroencephalogram. Anesth. Analg. 76, 976–8.

Seitzman, B.A., Abell, M., Bartley, S.C., Erickson, M.A., Bolbecker, A.R., Hetrick, W.P., 2016. Cognitive manipulation of brain electric microstates. Neuroimage 146, 0–1.

Seth, A.K., 2016. The real problem. Aeon https//aeon.co/.

Seth, A.K., Baars, B.J., 2005. Neural Darwinism and consciousness. Conscious. Cogn. 14, 140–68.

Seth, A.K., Barrett, A.B., Barnett, L., 2015. Granger Causality Analysis in Neuroscience and Neuroimaging. J. Neurosci. 35, 3293–3297.

Seth, A.K., Barrett, A.B., Barnett, L., 2011. Causal density and integrated information as measures of conscious level. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 369, 3748–3767.

Seth, A.K., Dienes, Z., Cleeremans, A., Overgaard, M., Pessoa, L., 2008. Measuring consciousness: relating behavioural and neurophysiological approaches. Trends Cogn. Sci. 12, 314–321.

Sewitch, D.E., 1984. The Perceptual Uncertainty of Having Slept: The Inability to Discriminate Electroencephalographic Sleep From Wakefulness. Psychophysiology 21, 243–259.

Shannon, C.E., 2001. A mathematical theory of communication. ACM SIGMOBILE Mob. Comput. Commun. Rev. 5, 3.

Siclari, F., LaRocque, J.J., Postle, B.R., Tononi, G., 2013. Assessing sleep consciousness within subjects using a serial awakening paradigm. Front. Psychol. 4, 1–9.

Silber, M.H., Ancoli-Israel, S., Bonnet, M.H., Chokroverty, S., Grigg-Damberger, M.M., Hirshkowitz, M., Kapen, S., Keenan, S. a., Kryger, M.H., Penzel, T., Pressman, M.R., Iber, C., 2007. The visual scoring of sleep in adults. J. Clin. Sleep Med. 3, 121–131.

Singer, W., Gray, C.M., 1995. Visual Feature Integration and the Temporal Correlation Hypothesis. Annu. Rev. Neurosci. 18, 555–586.

Sitt, J.D., King, J.-R., El Karoui, I., Rohaut, B., Faugeras, F., Gramfort, A., Cohen, L., Sigman, M., Dehaene, S., Naccache, L., 2014. Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. Brain.

Sokolov, D.D., Shvartsman, V.F., 1975. An estimate of the size of the universe from a topological point of view. Sov. Phys. 39, 196–200.

Soto, D., Silvanto, J., 2014. Reappraising the relationship between working memory and conscious awareness. Trends Cogn. Sci. 18, 520–525.

Spoormaker, V.I., Gleiser, P.M., Czisch, M., 2012. Frontoparietal connectivity and hierarchical structure of the brain's functional network during sleep. Front. Neurol. MAY, 1–10.

Sporns, O., 2010. Networks of the Brain. MIT Press, Cambridge, MA.

Sporns, O., Tononi, G., Edelman, G.M., 2002. Theoretical neuroanatomy and the connecti v ity of the cerebral cortex 135, 69–74.

Stam, C.J., Nolte, G., Daffertshofer, A., 2007. Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum. Brain Mapp. 28, 1178–1193.

Stamatakis, E.A., Adapa, R.M., Absalom, A.R., Menon, D.K., 2010. Changes in resting neural connectivity during propofol sedation. PLoS One 5.

Steriade, M., McCormick, D., Sejnowski, T., 1993a. Thalamocortical oscillations in the sleeping and aroused brain. Science (80-.). 262, 679–685.

Steriade, M., Nunez, A., Amzica, F., 1993b. Intracellular analysis of relations between the slow (< 1 Hz) neocortical oscillation and other sleep rhythms of the electroencephalogram. J. Neurosci. 13, 3266–3283.

Steriade, M., Nuñez, A., Amzica, F., 1993c. A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. J. Neurosci. 13, 3252–65.

Stevens, R.D., Hannawi, Y., 2016. Coma Prognostication: Looks That Count. Crit. Care Med. 44, 2292–2293.

Stevens, R.D., Sutter, R., 2013. Prognosis in Severe Brain Injury. Crit. Care Med. 41, 1104–1123.

Stickgold, R., Walker, M.P., 2007. Sleep-dependent memory consolidation and reconsolidation. Sleep Med. 8, 331–343.

Stoljar, D., 2016. The semantics of "What it's like" and the nature of consciousness. Mind 125, 1161–1198.

Storey, J.D., 2002. A direct approach approach to false discovery rates. J. R. Stat. Soc. 64, 479–498.

Strelets, V., Faber, P.L., Golikova, J., Novototsky-Vlasov, V., Koenig, T., Gianotti, L.R.R., Gruzelier, J.H., Lehmann, D., 2003. Chronic schizophrenics with positive symptomatology have shortened EEG microstate durations. Clin. Neurophysiol. 114, 2043–2051.

Šušmáková, K., Krakovská, a, 2007. Classification of waking, sleep onset and deep sleep by single measures. Meas. Sci. Rev. 7, 34–38.

Synek, V.M., 1988. Prognostically important EEG coma patterns in diffuse anoxic and traumatic encephalopathies in adults. J. Clin. Neurophysiol. 5, 161–74.

Tagliazucchi, E., Carhart-Harris, R., Leech, R., Nutt, D., Chialvo, D.R., 2014. Enhanced repertoire of brain dynamical states during the psychedelic experience. Hum. Brain Mapp. 35, 5442–5456.

Tanaka, H., Hayashi, M., Hori, T., 2000. Topographical characteristics of slow wave activities during the transition from wakefulness to sleep. Clin. Neurophysiol. 111, 417–427.

Tanaka, H., Hayashi, M., Hori, T., 1998. Topographic mapping of electroencephalography coherence in hypnagogic state. Psychiatry Clin. Neurosci. 52, 147–148.

Tanaka, H., Hayashi, M., Hori, T., 1997. Topographical characteristics and principal component structure of the hypnagogic EEG. Sleep 20, 523–534.

Taylor, N.R., 2013. Small World Network Strategies for Studying Protein Structures and Binding. Comput. Struct. Biotechnol. J. 5, e201302006.

Tegmark, M., 2016. Improved Measures of Integrated Information. PLOS Comput. Biol. 12, e1005123.

Téllez-Zenteno, J.F., Dhar, R., Wiebe, S., 2005. Long-term seizure outcomes following epilepsy surgery: A systematic review and meta-analysis. Brain 128, 1188–1198.

Thatcher, R.W., Cantor, D.S., McAlaster, R., Geisler, F., Krause, P., 1991. Comprehensive predictions of outcome in closed head-injured patients. The development of prognostic equations. Ann. N. Y. Acad. Sci. 620, 82–101.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., Farmer, J.D., 1992. Testing for nonlinarity in time series: the method of surrogate data. Phys. D 58, 77–94.

Tibshirani, R., 2011. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B (Statistical Methodol. 73, 273–282.

Tognoli, E., Kelso, J.A.S., 2014. The Metastable Brain. Neuron 81, 35–48.

Tomescu, M.I., Rihs, T. a., Becker, R., Britz, J., Custo, A., Grouiller, F., Schneider, M., Debbané, M., Eliez, S., Michel, C.M., 2014. Deviant dynamics of EEG resting state pattern in 22q11.2 deletion syndrome adolescents: A vulnerability marker of schizophrenia? Schizophr. Res. 157, 175–181.

Tononi, G., 2008. Consciousness as Integrated Information: a Provisional Manifesto. Biol. Bull. 215, 216–242.

Tononi, G., 2004. An information integration theory of consciousness. BMC Neurosci. 5, 42.

Tononi, G., Boly, M., Massimini, M., Koch, C., 2016. Integrated information theory: from consciousness to its physical substrate. Nat. Rev. Neurosci. 17, 450–461.

Tononi, G., Cirelli, C., 2006. Sleep function and synaptic homeostasis. Sleep Med. Rev. 10, 49–62.

Tononi, G., Edelman, G.M., 1998. Consciousness and complexity. Science 282, 1846–51.

Tononi, G., Koch, C., 2015. Consciousness: here, there and everywhere? Philos. Trans. R. Soc. B Biol. Sci. 370, 20140167–20140167.

Tononi, G., Massimini, M., 2008. Why does consciousness fade in early sleep? Ann. N. Y. Acad. Sci. 1129, 330–334.

Tsai, L.-L., Young, H.-Y., Hsieh, S., Lee, C.-S., 2005. Impairment of error monitoring following sleep deprivation. Sleep 28, 707–713.

Tsuchiya, N., 2017. "What is it like to be a bat?"—a pathway to the answer from the integrated information theory. Philos. Compass 12, 1–13.

Tsuchiya, N., Wilke, M., Frässle, S., Lamme, V.A.F., 2015. No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. Trends Cogn. Sci. 19, 757–770.

Tukey, J.W., 1949. Comparing individual means in the analysis of variance. Biometrics 5, 99–114.

Tye, M., 1992. Visual Qualia and Visual Content, in: Crane, T. (Ed.), The Contents of Experience. Cambridge University Press, pp. 158–76.

Uehara, T., Yamasaki, T., Okamoto, T., Koike, T., Kan, S., Miyauchi, S., Kira, J.I., Tobimatsu, S., 2014. Efficiency of a small-world brain network depends on consciousness level: A resting-state fMRI study. Cereb. Cortex 24, 1529–1539.

Van de Ville, D., Britz, J., Michel, C.M., 2010. EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. Proc. Natl. Acad. Sci. U. S. A. 107, 18179–18184.

Van Gaal, S., Lamme, V.A.F., 2012. Unconscious High-Level Information Processing. Neurosci. 18, 287–301.

Vanhaudenhuyse, A., Demertzi, A., Schabus, M., Noirhomme, Q., Bredart, S., Boly, M., Phillips, C., Soddu, A., Luxen, A., Moonen, G., Laureys, S., 2011. Two Distinct Neuronal Networks Mediate the Awareness of Environment and of Self. J. Cogn. Neurosci. 23, 570–578.

Varela, F., Lachaux, J.P., Rodriguez, E., Martinerie, J., 2001. The brainweb: Phase synchronization and large-scale integration. Nat. Rev. Neurosci. 2, 229–239.

Varela, F.J., 1996. Neurophenomenology A Methodological Remedy for the Hard Problem. J. Conscious. Stud. 3, 330–349.

Vassalli, A., Dijk, D.J., 2009. Sleep function: Current questions and new approaches. Eur. J. Neurosci. 29, 1830–1841.

Vert, J., Tsuda, K., Schölkopf, B., 2004. A primer on kernel methods. Kernel Methods Comput. Biol. 35–70.

Veselis, R.A., Feshchenko, V.A., Reinsel, R.A., Dnistrian, A.M., Beattie, B., Akhurst, T.J., 2004. Thiopental and Propofol Affect Different Regions of the Brain at Similar Pharmacologic Effects Robert. Anesth. Analg. 99, 912–918.

Vidaurre, D., Quinn, A.J., Baker, A.P., Dupret, D., Tejero-Cantero, A., Woolrich, M.W., 2016. Spectrally resolved fast transient brain states in electrophysiological data. Neuroimage 126, 81–95.

Vijayan, S., Ching, S., Purdon, P.L., Brown, E.N., Kopell, N.J., 2013. Thalamocortical mechanisms for the anteriorization of α rhythms during propofol-induced unconsciousness. J. Neurosci. 33, 11070–5.

Vinck, M., Oostenveld, R., van Wingerden, M., Battaglia, F., Pennartz, C.M.A., 2011. An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. Neuroimage 55, 1548–65.

Vos, P.E., Jacobs, B., Andriessen, T.M.J.., Lamers, K.J.B., Borm, G.F., Beems, T., Edwards, M., Rosmalen, C.F., Vissers, J.L.M., 2010. GFAP and S100B are biomarkers of traumatic brain injury. Neurology 75, 1786 LP-1793.

Vos, P.E., Lamers, K.J.B., Hendriks, J.C.M., van Haaren, M., Beems, T., Zimmerman, C., van Geel, W., de Reus, H., Biert, J., Verbeek, M.M., 2004. Glial and neuronal proteins in serum predict outcome after severe traumatic brain injury. Neurology 62, 1303 LP-1310.

Wang, J., Noh, G.J., Choi, B.M., Ku, S.W., Joo, P., Jung, W.S., Kim, S., Lee, H., 2017. Suppressed neural complexity during ketamine- and propofol-induced unconsciousness. Neurosci. Lett. 653, 320–325.

Watson, N.F., Badr, M.S., Belenky, G., Bliwise, D.L., Buxton, O.M., Buysse, D., Dinges, D.F., Gangwisch, J., Grandner, M.A., Kushida, C., Malhotra, R.K., Martin, J.L., Patel, S.R., Quan, S., Tasali, E., 2015. Recommended Amount of Sleep for a Healthy Adult: A Joint Consensus Statement of the American Academy of Sleep Medicine and Sleep Research Society. Sleep 38.

Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of “small-world” networks. Nature 393, 440–442.

Welch, 1984. A Technique for High-Performance Data Compression. Computer (Long. Beach. Calif). 17, 8–19.

Williams, D., 2018. Predictive Processing and the Representation Wars. Minds Mach. 28, 141–172.

Wilson, C., 2008. Up and down states. Scholarpedia 3, 1410.

Wright, K.P., Badia, P., Wauquier, A., 1995. Topographical and temporal patterns of brain activity during the transition from wakefulness to sleep. Sleep 18, 880–889.

Wu, D., Cai, G., Yuan, Y., Liu, L., Li, G., Song, W., Wang, M., 2011. Application of nonlinear dynamics analysis in assessing unconsciousness: A preliminary study. Clin. Neurophysiol. 122, 490–498.

Yuan, H., Zotev, V., Phillips, R., Drevets, W.C., Bodurka, J., 2012. Spatiotemporal dynamics of the brain at rest - Exploring EEG microstates as electrophysiological signatures of BOLD resting state networks. Neuroimage 60, 2062–2072.

Zahavi, D., 2004. Phenomenology and the project of naturalization. Phenomenol. Cogn. Sci. 3, 331–347.

Zhang, X.S., Roy, R.J., Jensen, E.W., 2001. EEG complexity as a measure of depth of anesthesia for patients. IEEE Trans. Biomed. Eng. 48, 1424–1433.

Zucco, G.M., Priftis, K., Stevenson, R.J., 2014. From blindsight to blindsmell: a mini review. Transl. Neurosci. 6, 8–12.